Genetic diversity of *Ophiocordyceps sinensis*, a medicinal fungus endemic to the Tibetan Plateau: Implications for its evolution and conservation

Yongjie Zhang†1,2, Lingling Xu†1,3, Shu Zhang2, Xingzhong Liu*1, Zhiqiang An4, Mu Wang5 and Yinglan Guo1

Address: 1Key Laboratory of Systematic Mycology and Lichenology, Institute of Microbiology, Chinese Academy of Sciences, No.3, 1st Beichen West Road, Chaoyang District, Beijing 100101, PR China, 2School of Life Sciences, Shanxi University, Wucheng Road, Xiaodian District, Taiyuan, Shanxi 030006, PR China, 3Graduate University, Chinese Academy of Sciences, Beijing,100049, PR China, 4Epitomics, Inc., 863 Mitten Road, Suite 103, Burlingame, CA 94010-1303, USA and 5School of Plant Science & Technology, Agriculture and Animal Husbandry College of Tibet, Nyingchi, Tibet 860000, PR China

Email: Yongjie Zhang - zhangyj2008@sxu.edu.cn; Lingling Xu - xull@im.ac.cn; Shu Zhang - zhangshu@sxu.edu.cn; Xingzhong Liu* - liuxz@im.ac.cn; Zhiqiang An - zhiqiang.an@epitomics.com; Mu Wang - wangmuth@yahoo.com.cn; Yinglan Guo - guoyl@im.ac.cn

* Corresponding author    †Equal contributors

Abstract

**Background:** *Ophiocordyceps sinensis* (syn. *Cordyceps sinensis*), endemic to alpine regions on the Tibetan plateau, is one of the most valuable medicinal fungi in the world. Huge commercial demand has led to excessive harvest and a dramatic decline in its numbers. The diversity of terrains and climates on the Tibetan Plateau and the broad insect host range (more than 50 species in the family Hepialidae) may have resulted in substantial intraspecific genetic diversity for this fungus. The objective of this study was to evaluate the population distribution of *O. sinensis* from geographically diverse regions of the Tibetan Plateau based on nrDNA ITS and *MAT1-2-1* gene sequences. Understanding of the genetic diversity and genesis of *O. sinensis* will provide important information for the evolution and conservation of this fungus.

**Results:** Significant sequence variations in the ITS and *MAT1-2-1* genes (27 and 23 informative sites, eight and seven haplotypes, respectively) were observed. Phylogenetic analysis based on ITS sequences, *MAT1-2-1* sequences, or their combined data set, clustered isolates from northern regions in one clade (clade I), whereas isolates from southern regions were dispersed in all four clades (clade I-IV). Single-strand conformation polymorphism (SSCP) analyses of 2639 ITS clones from seven samples revealed 91 different SSCP patterns that were subsequently sequenced. ITS heterogeneity was found in XZ-LZ07-H1 (Nyingchi population), and 17 informative sites and five haplotypes were detected from 15 clones. The five haplotypes clustered into three clades (clade I, II, and IV).

**Conclusions:** Significant genetic divergence in *O. sinensis* was observed and the genetic diversification was greater among southern isolates than that among northern isolates. The polymorphism of nrDNA ITS sequences suggested that *O. sinensis* spread from a center of origin (the Nyingchi District) to southern regions and subsequently to northern areas. These results...
Background
The ascomycete *Ophiocordyceps sinensis* (Berk.) Sung, Sung, Hywel-Jones and Spatafora (syn. *Cordyceps sinensis*) [1], commonly known as the Chinese caterpillar fungus, has been widely used in traditional Chinese medicine for the treatment of asthma, bronchial, lung inflammation, and other diseases [2]. The medicinal activity is from the complex of the moth caterpillar parasitized by the fungus and the fungal stroma biomass. The caterpillars, which belong to the family Hepialidae, live underground in soil burrows in the Tibetan plateau and feed on plant roots. When an infection is established, the fungus grows within body cavity of the insect larvae and eventually kills and mummifies them in the underground burrows. Fungal fruiting body emerges from the front end of the caterpillar and pokes out of the ground in spring and early summer [3]. The natural distribution of this fungus, however, is limited to alpine regions on the Tibetan Plateau and huge commercial demand has led to excessive harvest and a dramatic decline in its numbers [4]. As a result, the price for natural *O. sinensis* has increased rapidly and it cost as much as $32,000 (USD) per kg for top quality in late 2006 [5]. Although several fermented mycelial products from the anamorph (*Hirsutella sinensis*) of the fungus have been commercialized in China [6], the teleomorph, which is the component for traditional Chinese medicine, has not yet been commercially cultivated. The natural population of this fungus needs to be preserved for the fragile Tibetan Plateau ecosystem and for sustainable supply of this natural resource.

*Ophiocordyceps sinensis* occupies a diverse habitat on the Tibetan Plateau, which is the largest and highest plateau on Earth, covering more than 2.5 million km² at an average elevation of over 4.5 km [7]. It is surrounded by and interspersed with towering mountain ranges. Its complex geomorphology and climate generate substantial interregional variation which results in remarkable faunal and floral diversity and high levels of endemism [8-11]. *O. sinensis* is endemic to the Tibetan Plateau and occurs at elevations ranging from 3000 m up to the snow line [4]. This distribution pattern suggests that the evolution of *O. sinensis* was significantly influenced by the tectonic movements during the strong uplift of the Tibetan Plateau. The spread of *O. sinensis*, depending on the shooting of ascospores, is limited to defined areas in a mountain and is unlikely through discontinuous mountains. Therefore, gene flow between different populations on discontinuous mountains is expected to be low, and divergence may have occurred [12].

Intraspecific genetic diversity of *O. sinensis* may also be generated by its wide host range. *O. sinensis* infects soil-borne larvae of more than 50 species of ghost moths, mostly in the genus Hepialus and to a lesser extent in the genera Hepialiscus, Forkalus, and Bipectilus [13,14]. Coevolution between the fungus and its host insects may have originated a long time ago [15]. Although the host insects in general have a limited distribution in the Tibetan Plateau and their distribution varies among different mountain ranges and even from different sides and habitats of the same mountain [16], the greater mobility of moths may also drive the genetic diversity of the fungus.

To date, genetic diversity of *O. sinensis* has only been studied at the local level of limited geographic areas. Three populations found in the north (Menyuan, Maqu, and Luqu), the middle (Yushu and Chengduo), and the south (Baima Snow Mountain, Renzhi Snow Mountain, Dacaodi, and Chongcaoxiawa) of the Tibetan Plateau, have been recognized from 29 samples based on RAPD analyses [12]. Inter-simple sequence repeat (ISSR) analyses of *O. sinensis* from 11 counties in Qinghai Province also resulted in three groups (around the Qinghai Lake, central eastern Qinghai, and southern Qinghai) [17]. Based on ITS sequences of samples from 11 localities in southwestern China, *O. sinensis* was divided into two populations (one from southwestern Sichuan to the northern slopes of the Himalayan Mountains, the other from northern Sichuan to Qinghai and Gansu) [18]. However, ITS sequence analyses in another study did not indicate any subgroup for 17 *O. sinensis* samples from different geographical regions [19]. Overall, the number of *O. sinensis* samples used in earlier studies was limited, and more samples from Tibet, a major production area, should be surveyed. In addition, other molecular markers, especially protein-encoding genes, should be employed to clarify the genetic diversity and to refine the phylogenetic origin of this fungus.

Mating-type genes play an important role in the evolution of fungal species. Sexual development is controlled by a single mating-type locus (*MAT*) in the fungi of Ascomycota. The term "idiomorph" instead of "allele" is usually used to describe the two alternate forms (*MAT1-1* and *MAT1-2*) at the mating-type locus [20]. Mating-type genes evolve at a faster rate than other sequences, such as ITS
and glyceraldehyde 3-phosphate dehydrogenase [21]. Although MAT1-2-1 sequences have been used to investigate the phylogenetic relationships among closely related species and usually provide high resolution [22], other studies found that the variability of MAT1-2-1 within species is low [21]. A putative mating-type gene (MAT1-2-1) of O. sinensis was cloned in our laboratory. Similar to MAT1-2-1 orthologs of other ascomycetes, the O. sinensis MAT1-2-1 gene contained the conserved HMG motif. The sequence homologies to Cordyceps militaris and Cordyceps takaomontana were ~50% between cDNA sequences and ~40% between amino acid sequences (Zhang et al., unpublished data). During our preliminary experiment, several distinctive base changes within the MAT1-2-1 sequences were found among some O. sinensis isolates. Therefore, this gene was selected to investigate whether it can provide reasonable information on the genetic diversity of O. sinensis.

The objective of this study was to evaluate the population distribution of O. sinensis from geographically diverse regions of the Tibetan Plateau based on nrDNA ITS and MAT1-2-1 gene sequences. Understanding of the genetic diversity and genesis of O. sinensis should provide valuable information for the protection and sustainable utilization of this natural resource.

**Results**

**Sequence variations of nrDNA ITS and MAT1-2-1 genes among O. sinensis isolates**

Fifty-six O. sinensis isolates from different geographical regions were used in this study (Additional file 1). Their nrDNA ITS and MAT1-2-1 sequences were analyzed. For the 531-535 bp ITS sequences, the detectable nucleotide variations occurred at 35 sites, including 27 parsimony informative sites and eight singleton variable sites. Among the 27 parsimony informative sites, 13 were T/C base transitions; eight were A/G transitions; five were transversions of T/G, T/A, or A/C; and one was T/A/C substitution (see Additional file 2). Most of these informative sites were located in the ITS1 (11 sites) and ITS2 regions (14 sites), plus one each in the 5.8S and 28S regions (at sites 319 and 514, respectively). The singleton sites were not included in the subsequent haplotype analysis because one polymorphism at such a site was represented by only one isolate and such variation might be introduced artificially during PCR or sequencing steps. In addition, a 4-bp deletion was detected within the ITS2 region of two isolates (XZ-LZ07-H1 and XZ-LZ07-H2).

Detectable nucleotide variations within the 877-882 bp MAT1-2-1 sequences occurred at 28 sites, including 23 parsimony informative sites and five singleton variable sites. There were 10 T/C transitions, seven A/G transitions, and six C/G transversions (see Additional file 3). Nucleotide variations in nine sites resulted in amino acid changes. Interestingly, the two isolates that had base deletions within the ITS region contained a 5-bp insertion within the first intron of the MAT1-2-1 gene.

**Haplotypes within the 56 O. sinensis isolates**

Haplotype usually refers to a group of alleles at multiple loci that are transmitted together on a single chromosome or to a set of single nucleotide polymorphisms (SNPs) [23]. Ascosporic fungi are haploid and haplotypes were identified in this study according to base change patterns at informative sites among individual isolates. For the 56 isolates used in this study, eight ITS haplotypes, seven MAT1-2-1 haplotypes, and 14 combined haplotypes were detected (Figure 1, Additional file 1). The results showed that one haplotype might include isolates from different geographical regions (e.g., ITS type 7 and MAT1-2-1 type G), and isolates from the same location (e.g., XZ-NQ-176 and XZ-NQ-180) might contain different haplotypes.

For the eight ITS haplotypes (1-8), type 1 was the most dominant and it was represented by 28 isolates mainly from northern regions (Nagqu, Chamdo, Yushu, and Laji). Type 2 included nine isolates mainly from southern regions (Nyingchi and Baima). Other haplotypes were represented by two to five isolates (Figure 1, Additional file 1).

For the seven MAT1-2-1 haplotypes (A-G), type A was the most dominant and widespread; and it was represented by 33 isolates mainly from northern regions (Nagqu, Chamdo, Yushu, Laji, and Qilian). The next most common haplotype was type B, which included seven isolates from southern regions, such as Nyingchi, Baima, and Garze. Type C was represented by seven isolates from southern Tibet (Nyingchi and Mila). Other haplotypes were represented by one to three isolates (Figure 1, Additional file 1).

For the 14 combined haplotypes (Figure 1), type 1A was the most dominant and included 28 isolates. The next most common combined haplotype was type 2C, which was represented by six isolates from Nyingchi. Other haplotypes were represented by one to three isolates.

**Phylogenetic analyses**

Minimum evolutionary phylogenetic analyses were performed with sequences of the nrDNA ITS region, the MAT1-2-1 gene, and their combined data set. In these phylogenetic trees, two major clades (I and II) and two small clades (III and IV) were distinguished (Figure 1, see also Additional file 4).
**Figure 1**

Phylogenetic analyses of 56 *O. sinensis* isolates based on the combined data set of ITS and MAT1-2-1. Bootstrap values lower than 50% are not shown (the same hereinafter). Arabic numbers (1-8) indicate ITS haplotypes and capital letters (A-G) indicate MAT1-2-1 haplotypes. "N" or "S" behind the names of isolates represents northern or southern isolates, respectively.
Clade I (north clade) was composed of isolates mainly from northern Tibet (Nagqu and Chamdo) and Qinghai (Yushu, Laji, and Qilian) with the exception of several isolates from southern regions (XZ-ML-3, XZ-LZ07-64, and/or SC-3). This clade received bootstrap support values of 86%, 73%, and 95% in the three phylogenetic trees, and included three ITS haplotypes, two MAT1-2-1 haplotypes, and four combined haplotypes (Figure 1, Additional file 4).

Clade II was composed of southern isolates from southern Tibet (Nyingchi and Mila), Yunnan, and Sichuan. It had bootstrap support of 91%, 60%, and 64% in the three phylogenetic trees, and included three ITS haplotypes, three MAT1-2-1 haplotypes, and eight combined haplotypes (Figure 1, Additional file 4).

Two southern isolates, XZ-SN-44 and XZ-LZ05-6, clustered together as the third clade (clade III), having 97%, 99%, and 99% bootstrap support in the three trees (Figure 1, Additional file 4). The southern isolates XZ-LZ07-H1 and XZ-LZ07-H2 formed the fourth clade (clade IV) and had a distant relationship with other isolates (Figure 1, Additional file 4). Among all tested isolates, the maximal ITS distance value was 0.043 between XZ-LZ07-H1 and QH-QL-206 or QH-QL-207. The maximal MAT1-2-1 distance value was 0.023 between YN-4 and XZ-LZ07-H1 or XZ-LZ07-H2.

The phylogenetic trees constructed with ITS sequences and with MAT1-2-1 sequences had very similar topological structures (see Additional file 4). Four clades were discriminated in both trees, and composition of isolates for each clade was almost identical between the two trees. However, there were some discrepancies between the two trees. For example, isolate SC-3 clustered in clade II in the ITS tree, but in clade I in the MAT1-2-1 tree. Clade II was the sister group of clade III in the ITS tree, but the sister group of clade I in the MAT1-2-1 tree. The parameters of nucleotide diversity and haplotype diversity were usually higher in the ITS sequences than in the MAT1-2-1 sequences (Table 1). Therefore, the ITS sequence is relatively more polymorphic than the MAT1-2-1 sequence.

**Population structure and gene flow**

For both ITS and MAT1-2-1, significant genetic differentiation was detected among isolates in Nyingchi versus Yushu, Nagqu, Chamdo, and Garze; Garze versus Nagqu, Chamdo, and Nyingchi; and Baima versus Yushu, Nagqu, and Chamdo (Table 2). In addition, there was also significant differentiation for ITS sequences among isolates in Nagqu versus Qilian; Nyingchi versus Qilian; and Garze versus Yushu (Table 2). Analyses of molecular variance (AMOVA) based on haplotype frequencies showed significant $F_{ST}$ values in the following population grouping patterns: [south isolates] [north isolates]; [Nyingchi] [other south isolates] [north isolates]; [Garze] [other south isolates] [north isolates]; and [Garze] [other south isolates] Qilian and Laji [other north isolates] (Table 3). Significant difference ($P < 0.01$) was observed for these grouping patterns. These results indicate that an organized population structure exists among O. sinensis isolates.

### Table 1: Comparison of genetic diversity parameters

| Populations  | No. of sequences | No. of polymorphic sites | Average no. of nucleotide differences, $K$ | Nucleotide diversity, $Pi (×10^{-2})$ | Haplotype diversity, $h$ |
|--------------|------------------|--------------------------|------------------------------------------|------------------------------------------|--------------------------|
| Qilian (N)   | 2                | 0 (0)                    | 0.00 (0.00)                              | 0.00 (0.00)                              | 0.00 (0.00)              |
| Laji (N)     | 2                | 1 (0)                    | 1.00 (0.00)                              | 0.19 ± 0.09 (0.00)                       | 1.00 ± 0.50 (0.00)      |
| Yushu (N)    | 6                | 1 (1)                    | 0.33 (0.33)                              | 0.06 ± 0.04 (0.04 ± 0.02)                | 0.33 ± 0.22 (0.33 ± 0.22) |
| Nagqu (N)    | 12               | 2 (1)                    | 0.58 (0.17)                              | 0.11 ± 0.03 (0.02 ± 0.01)                | 0.53 ± 0.14 (0.16 ± 0.13) |
| Chamdo (N)   | 9                | 1 (0)                    | 0.22 (0.00)                              | 0.04 ± 0.03 (0.00)                       | 0.22 ± 0.17 (0.00)      |
| Nyingchi (S) | 11               | 22 (23)                  | 6.51 (7.42)                              | 1.25 ± 0.44 (0.85 ± 0.27)                | 0.62 ± 0.16 (0.80 ± 0.11) |
| Mila (S)     | 2                | 7 (4)                    | 7.00 (4.00)                              | 1.31 ± 0.66 (0.46 ± 0.23)                | 1.00 ± 0.50 (1.00 ± 0.50) |
| Garze (S)    | 5                | 2 (5)                    | 0.80 (2.00)                              | 0.15 ± 0.09 (0.23 ± 0.11)                | 0.40 ± 0.24 (0.70 ± 0.22) |
| Baima (S)    | 5                | 1 (2)                    | 0.60 (1.00)                              | 0.11 ± 0.03 (0.11 ± 0.03)                | 0.60 ± 0.18 (0.80 ± 0.16) |
| Northern regions | 31            | 10 (2)                   | 1.06 (0.13)                              | 0.20 ± 0.07 (0.02 ± 0.01)                | 0.49 ± 0.11 (0.13 ± 0.08) |
| Southern regions | 25            | 26 (27)                  | 4.37 (5.01)                              | 0.82 ± 0.26 (0.57 ± 0.17)                | 0.83 ± 0.06 (0.87 ± 0.04) |
| Overall      | 56               | 35 (28)                  | 4.45 (3.35)                              | 0.84 ± 0.13 (0.38 ± 0.09)                | 0.78 ± 0.05 (0.67 ± 0.06) |

* Fifty-six individuals were sequenced in this study (Additional file 1). "N" or "S" inside parentheses represents northern or southern populations, respectively (the same hereinafter). Numbers outside parentheses are values of ITS, and numbers inside parentheses are values of MAT1-2-1 (the same hereinafter). Genetic parameters of the Shannan population and the unknown population (Additional file 1) are included in the "southern regions" and "overall" statistics, but they are not given individually because they each possess one isolate. The "polymorphic sites" include both informative sites and singleton sites. For both ITS and MAT1-2-1, there is one polymorphic site shared by northern and southern isolates.
ITS heterogeneity by PCR-SSCP analyses

In total, 2639 ITS clones (1483 clones from four northern samples and 1156 clones from three southern samples) were analyzed by SSCP. Ninety-one clones with different SSCP patterns (60 for the northern samples and 31 for the southern samples) were sequenced (Table 4). Compared with the corresponding original sequence from direct sequencing, no ITS heterogeneity was detected for clones from each of these specimens except for XZ-NQ-166 and XZ-LZ07-H1, which had three (23 clones) and 17 informative sites (15 clones), respectively (Table 4). High nucleotide diversity (1.66 ± 0.34) and haplotype diversity (0.88 ± 0.07) were observed among clones of XZ-LZ07-H1. The 4-bp sequence deletion in XZ-LZ07-H1 from direct sequencing existed within some clones of XZ-LZ07-H1 (see Additional file 5). Among these 17 informative sites from 15 clones of XZ-LZ07-H1, 15 changes occurred at known informative sites and one at known singleton sites (see footnotes of Additional file 5). Every possible nucleotide change at these informative sites was shared by at least two to five clones. According to base changes at these informative sites, clones of XZ-LZ07-H1 can be divided into five haplotypes (I-V). The five haplotypes contained eight, three, two, one, and one clone(s), respectively (Fig-  

Table 2: Pairwise population differences and \( F_{ST} \) values for ITS and MAT1-2-I sequences

| Groups                        | Qilian (N) | Laji (N) | Yushu (N) | Nagqu (N) | Chamdo (N) | Nyingchi (S) | Mila (S) | Garze (S) | Baima (S) |
|-------------------------------|------------|----------|-----------|-----------|------------|--------------|----------|-----------|-----------|
| Qilian (N)                   | 0.00 (0.00)| 0.50 (0.00)| 0.73 (-0.30)| 0.56** (-0.33)| 0.81 (0.00)| 0.49** (0.31)| 0.50 (0.00)| 0.69 (0.48)| 0.55 (0.42)|
| Laji (N)                     | 0.50 (0.00)| 1.00 (0.00)| 0.09 (-0.30)| 0.01 (-0.33)| 0.27 (0.00)| 0.29 (0.31)| -0.33 (0.00)| 0.44 (0.48)| 0.28 (0.42)|
| Yushu (N)                    | 0.83 (0.00)| -0.08 (0.00)| 0.33 (0.33)| 0.01 (-0.03)| -0.06 (0.07)| 0.49** (0.34**)| 0.09 (0.09)| 0.64** (0.50)| 0.54** (0.45**)|
| Nagqu (N)                    | 0.73 (0.00)| -0.10 (0.00)| 0.01 (-0.01)| 0.53 (0.17)| 0.06 (-0.02)| 0.43** (0.48**)| 0.01 (0.38)| 0.52** (0.66**)| 0.44** (0.62**)|
| Chamdo (N)                   | 0.89 (0.00)| -0.06 (0.00)| -0.02 (0.00)| 0.03 (0.00)| 0.22 (0.00)| 0.56** (0.53**)| 0.62 (0.80)| 0.29 (-0.19)| 0.71** (0.75**)|
| Nyingchi (S)                 | 0.69 (0.51)| 0.19 (0.51)| 0.52 (0.36)| 0.43 (0.43)| 0.58 (0.51)| 0.56** (0.53**)| 0.62 (0.80)| 0.27 (0.66)| 0.46** (0.71**)|
| Mila (S)                     | 0.50 (0.00)| -0.25 (0.00)| -0.08 (-0.08)| -0.10 (-0.04)| -0.06 (0.00)| 0.19 (-0.17)| 1.00 (1.00)| 0.02 (0.21)| 0.28 (0.14)|
| Garze (S)                    | 0.80 (0.65)| 0.30 (0.65)| 0.63 (0.48)| 0.53 (0.57)| 0.69 (0.65)| 0.49 (0.20)| -0.10 (0.15)| 0.40 (0.70)| 0.43 (0.01)|
| Baima (S)                    | 0.70 (0.60)| 0.20 (0.60)| 0.53 (0.43)| 0.43 (0.52)| 0.59 (0.60)| 0.14 (0.16)| 0.20 (0.10)| 0.38 (0.01)| 0.60 (0.80)|

Notes: Above diagonal: pairwise \( F_{ST} \) values between populations. Diagonal elements: average number of pairwise differences within populations (PiX). Below diagonal: corrected average pairwise difference (PiXY-(PiX+PiY)/2). Pairwise \( F_{ST} \) values that are statistically different are indicated (\( P < 0.01 \)).

---

Table 3: AMOVA for grouping of populations estimated using F-statistics based on ITS and MAT1-2-I sequences

| Groups                        | Among pops within groups \( (F_{SC}) \) | Within pops \( (F_{ST}) \) | Among groups \( (F_{CT}) \) | Among groups \( (%) \) | Among groups \( (p) \) |
|-------------------------------|----------------------------------------|--------------------------|---------------------------|-----------------|----------------|
| [Nyingchi, Mila, Baima, Garze] [Qilian, Laji, Yushu, Nagqu, Chamdo] | 0.10 (0.02) | 0.24 (0.21) | 0.14** (0.21**| 28.60 (47.73)| 0.03 (<0.01)|
| [Nyingchi] [Mila, Baima, Garze, Qilian, Laji, Yushu, Nagqu, Chamdo] | 0.15 (0.11) | 0.24 (0.21) | 0.07 (0.08) | 15.82 (19.24)| 0.32 (0.32)|
| [Nyingchi] [Mila, Baima, Garze, Qilian, Laji, Yushu, Nagqu, Chamdo] | 0.08 (0.00) | 0.24 (0.21) | 0.14** (0.20**| 30.55 (49.85)| 0.02 (<0.01)|
| [Nyingchi, Mila, Baima] [Garze] [Qilian, Laji, Yushu, Nagqu, Chamdo] | 0.07 (0.00) | 0.24 (0.21) | 0.16** (0.21**| 33.62 (49.71)<0.01 (<0.01)|
| [Nyingchi, Mila, Baima] [Garze] [Qilian, Laji] [Yushu, Nagqu, Chamdo] | 0.05 (0.01) | 0.24 (0.21) | 0.17** (0.17**| 36.43 (44.40)<0.01 (<0.01)|

Notes: \( F_{SC} \) estimates the variation among populations relative to a regional grouping of populations. \( F_{ST} \) estimates the proportions of genetic variation within populations relative to the genetic variation for the whole samples. \( F_{CT} \) estimates the proportion of genetic variation among groups of populations relative to the whole species. \( F_{CT} \) values that are statistically different are indicated. * \( P < 0.05; ** P < 0.01. \)
Tables 4: PCR-SSCP results of four northern samples and three southern samples

| Specimens     | No. of tested clones | No. of sequenced clones | No. of polymorphic sites | Nucleotide diversity, P(×10^-2) |
|---------------|----------------------|-------------------------|--------------------------|--------------------------------|
|               | Stromata | Sclerotia | Mycelial vela | Total | Stromata | Sclerotia | Mycelial vela | Total |
| XZ-NQ-166 (N) | 285     | 247      | 145          | 677   | 4        | 8        | 11           | 23    | 3 (17)   | 0.37 ± 0.07 |
| QH-TS-197B (N)| 228     | 275      | 189          | 692   | 3        | 10       | 17           | 30    | 0 (18)   | 0.24 ± 0.05 |
| QH-TS-189 (N) | 36      | 16       | -            | 52    | 1        | 1        | -            | 2     | 0 (0)    | 0.00        |
| XZ-NQ-154 (N) | 40      | 22       | -            | 62    | 3        | 2        | -            | 5     | 0 (3)    | 0.23 ± 0.06 |
| XZ-LZ07-30 (S)| 255     | 234      | 16           | 505   | 3        | 0        | 4            | 7     | 0 (2)    | 0.11 ± 0.05 |
| XZ-LZ07-H1 (S)| 258     | 260      | 44           | 562   | 2        | 2        | 11           | 15    | 17 (10)  | 1.66 ± 0.34 |
| YN-6 (S)      | 38      | 51       | -            | 89    | 3        | 6        | -            | 9     | 0 (9)    | 0.38 ± 0.11 |

Total 1140 1105 394 2639 19 29 43 91 26 (49) 1.24 ± 0.15

* Two or three sections (stromata, sclerotia and/or external mycelial vela) were tested for each specimen.

Discussion

The distance value (K) value of nrDNA ITS sequences from one fungal species typically ranges from 0 to 0.05 [24,25]. The maximal distance value among O. sinensis populations in this study was 0.043. Although five putative new species related to O. sinensis were reported based on morphological characters and geographical distribution [26-28], they should not be treated as different species based on nrDNA ITS sequences [29]. A high degree of intraspecific genetic diversity of O. sinensis has been detected, but it was still highly dissimilar even from the most closely related lineages within same genus, having about 10% of ITS difference with O. robertsii (see Additional file 6).

In general, mating-type genes are more variable between species than within species [21] and the high degree of variation between species has been used to study the phylogeny of some closely related fungal groups [22]. In contrast to the low within-species variation (0.15-0.31% nucleotide sites different) of Cochliobolus heterostrophus [21] and Phaeosphaeria nodorum [30], our comparison of 56 O. sinensis isolates detected at least 23 base changes within 877-882 nucleotides of the MAT1-2-1 sequences (2.6%), including nine nonsynonymous variations. This suggests that the unique geophysical environment and wide insect host range have resulted in a high degree of intraspecific genetic diversity within O. sinensis. Mating-type genes are master regulators of the signal transduction pathway of mating [21], but whether variations within MAT1-2-1 sequence affect the sexual reproduction of O. sinensis is unknown. The fact that all strains contain the MAT1-2-1 gene suggests that it might be homothallic, capable of self-fertilizing. However, more study is needed to clarify the mating system of O. sinensis. This study is the first attempt to use a mating-type gene to determine whether the gene sequence differs among O. sinensis isolates from different localities. This is also the first protein-encoding gene used for the genetic analyses of O. sinensis isolates from different regions.

In this study, significant genetic divergence within the species of O. sinensis was revealed by two genes (nrDNA ITS and MAT1-2-1) and two main populations were detected: the southern population (Nyingchi, Mila, Shannan, Garze, and Baima) and the northern populations (Yushu, Qilian, Laji, Chamdo, and Nagqu). Diversification among southern isolates was greater than that among northern isolates. In addition, significant genetic differentiation was observed between Nyingchi and other populations. Nyingchi has greater within-population genetic diversity than other populations. ITS heterogeneity was detected in isolate XZ-LZ07-H1 from Nyingchi. These results imply that Nyingchi might have acted as a center of origin of O. sinensis and O. sinensis might have been transmitted from Nyingchi to other locations (Figure 3). O. sinensis might have spread eastward and westward first among southern regions on the Tibetan Plateau. This spread in southern Tibetan Plateau may be initiated along two large and parallel mountain ranges, the Himalaya Mountains (south of Nyingchi) and the Nyainqentanglha Mountains (north of Nyingchi), which run in east and west directions. The spread from south to north on the Tibetan Plateau might...
Phylogenetic analyses based on ITS sequences of 91 clones and their original sequences from direct sequencing. *•* indicates clones of XZ-LZ07-H1. Names of clones beginning with 154, 166, 189, and 197 originate from northern isolates XZ-NQ-154, XZ-NQ-166, QH-YS-189, and QH-YS-197B, respectively. Names of clones beginning with 6, 30, and H originate from southern isolates YN-6, XZ-LZ07-30, and XZ-LZ07-H1, respectively. Ca, Ch, and T indicate the source of clones from stromata, sclerotia, and external mycelial vela, respectively.
have occurred at a later time and this may be achieved by air currents from the Indian Ocean through Yarlung Zangbo Grand Canyon.

Both *O. sinensis* and its host insects are endemic species on the Tibetan Plateau; and the geography and climate may have played an important role in their evolution and coevolution. However, the broad range of host insects and their distribution characters may also have impacted on the diversification of *O. sinensis* because the fungus must parasitize caterpillar to complete its life cycle [13,14,16]. Host species of the fungal isolates used in this study and the genetic diversity of the hosts were not determined. The whole picture of *O. sinensis* genetic differentiation and the coevolution between *O. sinensis* and host insects needs to be further studied [15].

Two southern isolates, XZ-LZ07-H1 and XZ-LZ07-H2, had relatively distant genetic relationships from other isolates. Compared to the ITS or *MAT1-2-1* sequences of other isolates, the sequences of these two isolates had two fixed 4- or 5-bp indels and 13 unique base changes. According to the newly proposed "indel-associated substitution" theory, the high substitution frequency in these two isolates may be related to the existence of the two indels. The theory asserts that the single-nucleotide mutation rate increases in regions surrounding the insertion or deletion sites [31]. Furthermore, ITS heterogeneity was detected in XZ-LZ07-H1, but the potential sources of heterogeneity are still unknown. The fact that one clone of XZ-LZ07-H1 clustered in the northern clade suggests that hybridization is one of the possible mechanisms. Inherent mechanisms involving slippage events during DNA duplication may

---

**Figure 3**

Proposed center of origin and transmission pathway of *O. sinensis* on the Tibetan Plateau. The figure at the upper left corner is the enlargement of Nyingchi District (proposed center of origin). Haplotype of *O. sinensis* isolates at each sampling site are shown on the map. Original map was scanned from Liao, 1990 [41].
also be one of the reasons for the existence of base indels in the sequence of XZ-LZ07-H1 and XZ-LZ07-H2.

The considerable base changes detected in this study indicated that recombination might have occurred between different O. sinensis populations. Considering the difficulty of gene flow among different populations, O. sinensis may have an ancestral origin. According to a recent report, although diversification of insect pathogens within the fungal order of Hypocreales occurred during the Cretaceous period, the genesis of O. sinensis was an event during the Cenozoic era (65.5 million years ago (Mya) to present) [32]. We suggest the history of O. sinensis was not earlier than 1.8 Mya, because the altitude of the Tibetan Plateau reached ~2000 m in the Eopleistocene (1.8 to 0.8 Mya) and it has kept a trend of rapid uplifting since then [33].

Conclusions

Based on two DNA sequences (nrDNA ITS and MAT1-2-1), significant genetic divergence was detected among 56 isolates of O. sinensis collected from different regions of the Tibetan Plateau. The genetic diversification was greater among southern isolates than that among northern isolates, indicating that southern populations were important genetic reservoirs of this species to be considered in conservation programs. The polymorphism of nrDNA ITS sequences suggests that the Nyingchi District is the center of origin of O. sinensis; and that the fungus may have first spread from this center first to other southern regions and then to northern areas.

Methods

Ophiocordyceps sinensis

Fifty-six natural specimens representing 11 populations of O. sinensis were collected from the Tibet Autonomous Region, Qinghai Province, Sichuan Province, and Yunnan Province in southwestern China during 2005 and 2008 (Additional file 1). After dig out from the soil, each specimen was placed into a new zip plastic bag immediately. They were taken back to our laboratory by keeping at low temperature in a portable refrigerator and stored at 4°C after arrival. Anamorphic strains were isolated from stromal tissues within three weeks according to the protocol of Liu et al [34]. The remaining fruiting bodies were stored at -20°C. Either axenic isolates or the stromatal section of natural specimens were used for DNA analysis.

Amplification of sequences from the nrDNA ITS region

Genomic DNA was extracted with the CTAB method [35]. ITS regions of nrDNA were amplified with the primers ITS5 (5'-GGAAGTTAAGATCTGAAACGAGG-3') and ITS4 (5'-TCCTCCGGCTATTGATATGC-3') using high-fidelity Pfu DNA polymerase (TIANGEN BIOTECH, China). The PCR mixture contained 5 μl of 10× Pfu buffer, 5 μl of dNTP mixture (2.5 mM), 2 μl of each primer (10 μM), 2 μl of deionized formamide, 1 μl of MgCl2 (25 mM), 1 μl of genomic DNA, and 0.5 μl of Pfu DNA polymerase in a total volume of 50 μl. PCR amplification was performed in a TGradient™ thermocycler (Biomетra, Germany). The mixture was heated for 5 min at 94°C; then subjected to 35 cycles of 40 s at 94°C, 40 s at 54°C, and 60 s at 72°C; and a final 10 min elongation step at 72°C. After confirmation of the PCR products by agarose gel electrophoresis, the fragments were cleaned using the 3S Spin PCR Product Purification Kit (Biocolor Bioscience & Technology Company, China). The purified PCR products were then sequenced with the primer ITS5 using an ABI 3730 XL DNA sequencer with BigDye 3.1 Terminators (Applied Biosystems, USA).

Amplification of sequences for MAT1-2-1 gene

The MAT1-2-1 gene of O. sinensis was first cloned in our laboratory using degenerate primers, and both flanks were extended using DNA Walking SpeedUp™ Premix Kit (Seegene, Korea) (Zhang et al., unpublished data). For this study, MAT1-2-1 sequences were amplified with the primers MAT1-2F1 (5'-CCACCGATTTCAAGTCTCTT-3') and MAT1-2R2 (5'-CATTTTCAGTCGCTGTCGTG-3'). The assay included: 5 μl of 10× Pfu buffer, 5 μl of dNTP mixture (2.5 mM), 2 μl of each primer (10 μM), 1 μl of genomic DNA, 0.5 μl of Pfu DNA polymerase, and water to a final volume of 50 μl. The PCR profile consisted of: 94°C for 5 min; followed by 35 cycles of 94°C for 40 s, 55°C for 40 s, and 72°C for 70 s; with a final 10 min incubation at 72°C. PCR products were purified, and sequencing was done in both directions using both the forward and reverse primers following the protocol described above.

SSCP analysis

SSCP is the electrophoretic separation of single-stranded nucleic acids based on subtle differences in sequence (often a single base pair). We used this technology to identify ITS heterogeneity of O. sinensis. Based on the results from analyses of ITS sequences, four samples from northern regions (QH-YS-197, QH-YS-189, XZ-NQ-154, and XZ-NQ-166) and three samples from southern regions (XZ-LZ07-30, XZ-LZ07-H1, and YN-6) were selected for SSCP analyses. Each sample was divided into two or three sections (stromata, sclerotia, and/or external mycelial vela), and genomic DNA was extracted separately from each section. ITS sequences were amplified with the primers ITS1 (5'-TCCTCGTAGTAGTGAACCTGCGC-3') and ITS4 using the same conditions described for amplification of nrDNA ITS sequences. The exception was that Taq DNA polymerase (TaKaRa, Japan) instead of Pfu was used for the convenience of subsequent cloning. The PCR products were ligated into the TA plasmid pMD18-T (TaKaRa, Japan) and transformed into Escherichia coli DH5α. Col-
ony PCR was then conducted to amplify the ITS1 region with the primers ITS1 and ITS2 (5’-GCTGCCTCTTATGC-GATGC-3’). Clones with expected amplicons were used for SSCP analyses following the procedure of Wang et al [36]. Different migration profiles between clones were compared using the Image J software. Representative clones of distinct patterns were sequenced with the universal primer M13-47.

**Sequence analyses**

Sequences were aligned with Clustal X [37], and ambiguous regions in both sides were excluded from the subsequent analyses. For the ITS region, 531-535 bp of sequences including partial 18S (5 bp), ITS1 (160 bp), 5.8S (157 bp), ITS2 (171-175 bp), and partial 28S (38 bp) were used. For the MAT1-2-1 gene, 877-882 bp of sequences from 14 bases upstream of the start codon to six bases downstream of the stop codon were used. Pairwise distance matrices and minimum evolutionary (ME) phylogenetic analyses were conducted with the Kimura 2-Parameter model using MEGA 4 software [38]. To assess the confidence of phylogenetic relationships, the bootstrapping method was performed using MEGA 4 (1000 resamplings). Bootstrap values lower than 50% are not shown.

**Additional file 2**

*Sequence differences among eight haplotypes found in the nrDNA ITS regiona.* "NO" indicates one base deletion for haplotype 8 at position of 492.

Click here for file [http://www.biomedcentral.com/content-supplementary/1471-2148-9-290-S2.XLS]

**Additional file 3**

*Sequence differences among seven haplotypes found in the MAT1-2-1 sequencesb.*

Click here for file [http://www.biomedcentral.com/content-supplementary/1471-2148-9-290-S3.XLS]

**Additional file 4**

*Phylogenetic analyses of 56 O. sinensis isolates based on ITS sequences (A) and MAT1-2-1 sequences (B).* Bootstrap values lower than 50% are not shown.

Click here for file [http://www.biomedcentral.com/content-supplementary/1471-2148-9-290-S4.TIFF]

**Additional file 5**

*Alignment of ITS sequences from direct sequencing and from clones of XZ-L207-H1a.* Description: Dashes and dots indicate identity and indel, respectively. HCa, HCh, and HT represent clones from stromata, sclerotia, and external mycelial velata, respectively. From these clones, 17 informative sites were detected, of which 15 occurred at previously known informative sites (6, 7, 13, 25, 30, 115, 132, 350, 379, 395, 436, 442, 446, 463, 492), one at a previously known singleton site (32).

Click here for file [http://www.biomedcentral.com/content-supplementary/1471-2148-9-290-S5.TIFF]

**Additional file 6**

*Simplicity of ITS sequences for O. sinensis and other taxa.*

Description: * These taxa were chosen according to Sung et al (2007).*

Click here for file [http://www.biomedcentral.com/content-supplementary/1471-2148-9-290-S6.XLS]

**Acknowledgements**

This study was supported by grants from National Plans of Science and Technology (Grant No. 2007BAI32B04), National Outstanding Youth Foundation (30625001), and National 863 Plan of China (2007AA021506). The authors are grateful to Prof. Zhang Luo, Dr. Bingda Sun and Miss Yang Hao for their help in collecting samples. We acknowledge Dr. Ence Yang for his advice in analyzing sequences. We also thank Prof. Bruce Jaffee (the University of California at Davis) for serving as pre-submission reviewers and for his valuable comments and suggestions.

**References**

1. Sung GH, Hywel-Jones NL, Sung M, Luangsa-ard JJ, Shrestha B, Spatafora JW. Phylogenetic classification of Cordyceps and the clavicipitaceous fungi. Stud Mycol 2007, 57(1):5-59.
2. Zhu JS, Halpern GM, Jones K: The scientific rediscovery of an ancient Chinese herbal medicine: Cordyceps sinensis part I. Altern Complement Med 1998, 289-303.
6. Ke CK, Cong ZY, Kang SC, Smirnov A, Holben B: BMC Evolutionary Biology 2008, 21:7.
7. Winkler D: Journal of Evolutionary Biology 1999, 22:143-148.
8. Winkler D, Yartsa Gunbu (Cordyceps sinensis) and the fungal transmission: Journal of Evolutionary Biology 2008, 21:291-305.

5. Turgeon BG: PLoS Biology 2008, 6(2):42-49.
6. Li JJ, Fang XM: Uplift of the Tibetan Plateau and environmental changes. Chin Sci Bull 1999, 44(23):2117-2124.
7. Tapponnier P, Xu ZQ, Roger F, Meyer B, Arnaud N, Wittlinger G, Yang JS: Identification and differences of yeasts species by single-strand conformation polymorphism of ribosomal DNA. Appl Environ Microbiol 2008, 74(9):2604-2611.
8. Tian DC, Wang Q, Zhang PF, Araki H, Yang SH, Kreitman M, Nagylaki T, Hudson R, Bergelson J, Chen JQ: Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. Nature 2008, 455:105-108.
9. Doyle JJ, Doyle JL: A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bulletin 1987, 19:11-15.
10. Zhang FF, Jiang ZG: Morphological and phytogenetic studies on the genus Hepialus and their geographical distribution. Acta Entomol Sin 1996, 39(4):413-422.
11. Zhang FF, Jiang ZG: Mitochondrial phylogeny and genetic diversity of Tibetan gazzoles (Procorea picticaudata): Implications for conservation. Mol Mycol 2006, 41(2):313-321.
12. Chen YJ, Zhang YP, Yang YX, Yang DR: Genetic diversity and taxonomic implication of Cordyceps sinensis as revealed by RAPD markers. Biochim Biophys Acta 1999, 14:5-6:201-213.
13. Liu F, Wu XL, Yin DH, Chen SJ, Zeng W: Overview in biological studies of host insects of Cordyceps sinensis. Chongqing J Res Chin Drugs 2005, 45:52.
14. Zhu HF, Wang LY, Han XH: Fauna Sinica: Insecta, Lepidoptera Hepialoae Epiplemidae Volume 38. Beijing: Science Press, 2004.
15. Chen YJ, Yang DR, Yang YX, Zhang YP: Molecular evolution of Cordyceps sinensis from seven regions on the Tibetan Plateau. In Proceedings of the Third Youth Annual Conference of the China Association for Science and Technology (Materials Science and Engineering):20 August 1998; Beijing: China Science and Technology Press; 1998:89.
16. Yang DR, Li CD, Shu C, Yang YX: Studies on Chinese species of the genus Hepialus and their geographical distribution. Acta Entomol Sin 1996, 39(4):413-422.
17. Li H, Cheng Z, Yang XL, Li S, Zhou TS, Zhang WJ, Chen JK: Genetic variation and affinity of Cordyceps sinensis in Qinghai Province based on analysis of morphologic characters and inter-simple sequence repeat markers. Chin Tradit Herb Drugs 2005, 36(12):1859-1864.
18. Kinjo N, Zhang M: Morphological and phylogenetic studies on Cordyceps sinensis distributed in southwestern China. Myco- science 2001, 42(6):567-574.
19. Kinjo N, Zhang M: Notes on the alpine Cordyceps of China and nearby nations. Mycoscience 1998, 39(2):157-167.
20. Turgeon BG, Yoder OC: Proposed nomenclature for mating type genes of filamentous ascomycetes. Fungal Genet Biol 2000, 31(1):1-5.
21. Turgeon BG: Application of mating type gene technology to problems in fungal biology. Annu Rev Phytopathol 1998, 36(1):115-137.
22. Groenewald M, Groenewald JZ, Harrington TC: Mating type gene analysis in apparently asexual Cercopis species is suggestive of cryptic sex. J Genet Biol 2006, 43(12):813-825.
23. Silverman EK: A molecular phylogeny approach to the genus Cercopis. Proc Am Thorac Soc 2007, 4:4-8.
24. Qu LH, Chen YQ: Key to molecular taxonomy of Cercopis species. Acta Sci Nat Univ Sunyatseni 1999, 38(1):1-7.
25. Hibbet DS: Ribosomal RNA and fungal systematics. Trans Mycol Soc Japan 1992, 33:533-556.
26. Zhang M, Kinjo N: Notes on the alpine Cordyceps of China and nearby nations. Mycoscience 1998, 39(2):16-22.
27. Zang M, Yang DR, Li CD: A new taxon in the genus Cordyceps from China. Mycologia 1990, 72:387-392.
28. Zhang KY, Wang CJ, Yan MS: A new species of Cordyceps from Gansu, China. Trans Mycol Soc 1989, 30:295-299.
29. Jiang Y, Yao Y: Current understanding of molecular systematics of Cordyceps. J Fungal Res 2004, 2(1):56-67.
30. Bennett RS, Yun SH, Lee TY, Turgeon BG, Arseniuk E, Cunfer BM, Bergstrom GC: Identity and conservation of mating type genes in geographically diverse isolates of Phaeo sperma nodorum. Fungal Genet Biol 2001, 40(1):25-37.
31. Tian DC, Wang Q, Zhang PF, Araki H, Yang SH, Kreitman M, Nagylaki T, Hudson R, Bergelson J, Chen JQ: Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. Nature 2008, 455:105-108.
32. Sung GH, Poinar GO Jr, Spatafora JW: The oldest fossil evidence of animal parasitism by fungi supports a Cretaceous diversification of fungal-arthropod symbioses. Mol Phylogenet Evol 2008, 49(2):495-502.
33. Zhang DF, Li FQ, Bian JM: Environmental effects of the Qinghai-Tibetan Plateau uplift during the Quaternary in China. Environ Geol 2000, 39(12):1352-1358.
34. Liu XJ, Guo YL, Yu YX, Zeng W: Isolation and identification of the anamorphic state of Cordyceps sinensis (Berk.) Sacc. Acta Mycol Sin 1989, 8:35-40.
35. Doyle JJ, Doyle JL: A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bulletin 1987, 19:11-15.
36. Wang QQ, Wang SA, Bai FY: Rapid differentiation of phenotypically similar yeast species by single-strand conformation polymorphism of ribosomal DNA. Appl Environ Microbiol 2008, 74(9):2604-2611.
37. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 1997, 25:487-490.
38. Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA software version 4.0. Mol Biol Evol 2007, 24:1596-1599.
39. Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R: DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 2003, 19:2496-2497.
40. Laval Excoffier LG, Schneider S: Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 2005, 1:47-50.
41. Liu K: The Atlas of the Tibetan Plateau Beijing: Science Press; 1990.