Article

Chinese Science Bulletin

Genetics

August 2013 Vol.58 No.23: 2846–2854
doi: 10.1007/s11434-013-5929-5

Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus

HU Xiao[1†], ZHANG YongJie[2†], XIAO GuoHua[1†], ZHENG Peng[1], XIA YongLiang[1], ZHANG XingYu[3], ST LEGER Raymond J[4], LIU XingZhong[3] & WANG ChengShu[1*]

[1] *Key Laboratory of Insect Developmental and Evolutionary Biology, Institute of Plant physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China;*
[2] *School of Life Sciences, Shanxi University, Taiyuan 030006, China;*
[3] *State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;*
[4] *Department of Entomology, University of Maryland, College Park, MD 20742, USA*

Received February 24, 2013; accepted May 21, 2013; published online June 17, 2013

The caterpillar fungus *Ophiocordyceps sinensis* (best known as *Cordyceps sinensis*) mummifies ghost moth larvae exclusively in Tibetan Plateau alpine ecosystems. Touted as "Himalayan Viagra", the fungus is highly prized due to its medical benefits and dwindling supplies. Attempts to culture the sexual fruiting-body have failed and the huge market demand has led to severe devastation of local ecosystems and to the fungus heading towards extinction. By genome sequencing, we establish that unlike related insect pathogens *O. sinensis* contains two compatible mating-type genes in its genome and is self-fertile, i.e. homothallic. However, sexual processes are only initiated under native environmental conditions. *O. sinensis* resembles biotrophic plant pathogens in having a genome shaped by retrotransposon-driven expansions. The resulting changes in gene content suggest that *O. sinensis* has a biphasic pathogenic mechanism beginning with stealth pathogenesis in early host instars. *O. sinensis* is the first psychrophilic fungus sequenced and is adapted to extreme cold with putative antifreeze proteins and mechanisms for increasing lipid accumulation and fatty acid unsaturation. We hypothesize that for the inbreeding *O. sinensis* the massive proliferation of retrotransposons provides a tradeoff between the advantages of increased genetic variation independent of sexual recombination and deletion of genes dispensable for its specialized pathogenic lifestyle.

***Ophiocordyceps sinensis*, genome expansion, homothallism, biotrophic parasitism, psychrophile**

Citation: Hu X, Zhang Y J, Xiao G H, et al. Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus. Chin Sci Bull, 2013, 58: 2846–2854, doi: 10.1007/s11434-013-5929-5

The caterpillar fungus *Ophiocordyceps sinensis* (best known as *Cordyceps sinensis*; anamorph: *Hirsutella sinensis*) is one of the most highly valued traditional Chinese medicines [1,2]. First documented in the 15th century [2], the fungus is only found at high altitudes (>4000 m) on the Tibetan Plateau. Temperatures here drop to less than –40°C in winter and *O. sinensis* is a psychrophile with an optimum growth temperature at 18°C [3]. The fungus colonizes ghost moth caterpillars (*Thitarodes* spp.) forming a parasitic complex that comprises the remains of the caterpillar and fungal sexual stroma (Figure 1(a)). Many studies have demonstrated the diverse bioactive ingredients and broad medical effects of *O. sinensis* [4,5]. Touted as "Himalayan Viagra", the parasitic complex fetches $60000–$75000 per kilogram and is used to treat a variety of ailments including impotence, fatigue and cancer [5,6]. *O. sinensis* grows slowly on artificial media (Figure 1(b) and (c)) but attempts at cultivating the fungus to produce fruiting bodies have consistently failed [6,7]. The huge market demand has led to over harvesting (Figure 1(d)), severe devastation of fragile alpine environments (Figure 1(e)), and multiple homicides as villagers tried to prevent outsiders from cashing in on their limited supply [5]. The parasitic complexes are hand-

†These authors contributed equally to this work.
*Corresponding author (email: cswang@sibs.ac.cn)

Figure 1 Phenotyping and field survey of *O. sinensis*. (a) A fruiting body of *O. sinensis* growing out of the head of a mummified ghost moth caterpillar in its natural habitat of the alpine Tibetan Plateau (Inset, thread-like ascospores discharged from mature fruiting body). (b) *O. sinensis* colony grown on a solid nutrient-rich medium for 70 d at 18°C (Inset, pear-shaped conidia). (c) Bar-shaped blastospores produced by *O. sinensis* grown in a liquid culture for one month (Arrow points to the contraction ring required for yeast-like budding). Bar, 5 μm. (d) Collection of parasitic complexes by local villagers at Seji Lashan, Nyingchi County, Tibet (photo taken on May 20, 2006). (e) Devastation of alpine environments due to specimen collection (digging holes are arrowed) at Biru County, Nagqu District, Tibet (photo taken on June 12, 2007).

collected in the field (Figure 1(d)) and immature specimens that have not yet discharged reproductive spores are preferred as they fetch a higher market price [6]. This further reduces the ability of *O. sinensis* to replenish its population and the fungus seems to be heading towards extinction in nature [5].

The biology of *O. sinensis* remains mysterious largely owing to difficulties in observing fungal development and infection processes in the field. The route of infection is unknown but probably occurs at the first instar larval stage [8]. The host caterpillars live underground for 4–5 years and have 7–9 instars depending on their sex [9]. During most of this time the fungus is believed to remain dormant and is only observed *in insecta* in later instars just preceding the host's death. The fungus then fully colonizes the cadaver and produces a sexual structure after over wintering, the so-called "winter worm, summer grass" in Chinese literally [1,9]. The molecular basis for this lifestyle is entirely unknown as is the sex mode of *O. sinensis*.

Sexual identity varies among different fungal species. Alternative genes, e.g. either *MAT1-1* or *MAT1-2*, at the mating type (Mat) locus control transitions between heterothallism, homothallism and pseudohomothallism [10,11]. To illuminate *O. sinensis* sexuality and the molecular basis of its lifestyle we sequenced the genome of *O. sinensis* and found that, in contrast to its close relatives, *O. sinensis* contains two compatible Mat loci in the genome and is sexually self-fertile. Genomic information and infection studies suggest the fungus employs a mechanism of stealth pathogenesis. *O. sinensis* is a psychrophilic fungus. We found putative

adaptations to extreme cold in the form of increased lipid accumulation and fatty acid unsaturation, and novel antifreeze proteins.

1 Materials and methods

1.1 Fungal strains and maintenance

O. sinensis strain Co18 (CGMCC 3.14243, cataloged at the Center of General Microorganisms Culture Collection, China) was selected for genome sequencing. The strain was purified from the stromal tissue of a fruiting-body sample originally collected from the Yushu area, Qinhai Province. Natural specimens were collected and placed in a portable refrigerator, brought back to the lab and stored at 4°C. Multiple single ascospore isolates were prepared as described previously [12] and used for genotyping the mating-type loci (Table S1). Fungal cultures were maintained either in liquid (2% sucrose, 0.5% mannose, 0.15% galactose, 0.35% tryptone, 1.5% yeast extract, 0.15% $MgSO_4$, 0.3% KH_2PO_4) or on solid (liquid medium amended with 2% agar) media at 18°C as previously described [3,12]. Fungal cells harvested from the solid or liquid medium were stained with the fluorescent dye Bodipy (Invitrogen) as previously described [13].

1.2 Genome sequencing and assembly

Genomic DNA of Co18 was extracted from the liquid culture for shotgun sequencing with a Roche 454 GS FLX system. This resulted in 223 Mb of sequence data with an average read length of 400 bp. Assembly was performed using the Newbler software (Ver. 2.3) installed within the Roche 454 suite package, which produced 25439 contigs. Sequencing and scaffolding were also performed using an Illumina Hiseq 2000 platform. DNA libraries of 200 bp, 500 bp and 2 kb inserts were generated and sequenced. This resulted in 10.6 Gb of paired-end and mate-pair reads. To avoid adaptor contamination and exclude low-quality reads and artificial duplication, a series of strict filtering steps were performed before assembly using a SOAP denovo program (ver. 1.05) [14]. The genome size was estimated using the total length of sequence reads divided by sequencing depth by counting the copy number of a certain k-mer (e.g. 15-mer) present in sequence reads. From this we estimated the *O. sinensis* genome size to be ~120 Mb. The actual assembled size was 87.7 Mb. To assess the completeness of gene catalogs in the draft genome, a core eukaryotic genes mapping approach (CEGMA) was used for analysis [15]. The whole genome sequence has been deposited at DDBJ/EMBL/GenBank under the accession No.: ANOV00000000.

1.3 Repeat rate and pseudogene analysis

For overall repeat ratio analysis, the assembled contigs from the Illumina paired-end libraries (200 and 500 bp) were classified by length into short contigs (⩽ 100 bp) and long contigs (>100 bp). Repeat contigs were defined as that the short contigs which sequencing depth are more than 1.5-fold average coverage and the long repeat contigs which depth are more than 1.8-fold average coverage. Relative to the total contigs and contig length, ratios of repeated contig numbers and lengths were then calculated. The assembly was also masked for different types of repeats with the program RepeatMasker (http://www.repeatmasker.org/) (Ver. open-3.3.0). Genome mapping of repeats was conducted using the program OmniMapFree (http://www.omnimapfree.org/). Pseudogene identification was conducted with the PseudoPipe program with default settings [16].

1.4 Annotation and protein family analysis

Gene prediction was conducted as previously described with additional manual annotation checks [17–19]. Whole genome protein families were classified by InterproScan analysis (http://www.ebi.ac.uk/Tools/pfa/iprscan/). Protein family analyses of proteases, cytochrome P450s (CYPs), carbohydrate-active enzymes and putative virulence-related proteins were performed as we previously described [17]. Briefly, the families of proteases were identified by Blastp searching against the MEROPS peptidase database Release 9.4 with a cut-off of *E* value of 1e-20 (http://merops.sanger.ac.uk/). CYP proteins were named according to the classifications collected at the P450 database (http://blast.uthsc.edu/). Carbohydrate-active enzymes were classified based on the library of catalytic and carbohydrate-binding module enzymes (http://www.cazy.org/). Dehydrogenases were classified based on the curated enzymes cataloged at the DBD database (http://www.bifku.in/DBD/). Putative virulence factors of *O. sinensis* were identified by searching against the pathogen-host interaction database (http://www.phi-base.org/about.php) with a cut-off *E* value of 1×10^{-5}. Fungal sex-related genes that have been functionally verified in the model ascomycetes *Aspergillus nidulans* and *Neurospora crassa* [20] were used for blastp search against the genomes of *O. sinensis* and other insect pathogens to retrieve the respective homologs. To identify the gene clusters and their proteins responsible for the biosynthesis of secondary metabolites, the whole genome dataset was subject to analyses with the programs antiSMASH [21] and SMURF [22]. To determine putative antifreeze proteins (AFPs), whole genome analysis was performed for *O. sinensis* and other insect pathogens for different type of AFPs, e.g. alanine-rich (> 60%), molecular weight (MW) 3.4–4.5 kD and forming α-helix in secondary structure for Type I AFPs; cysteine-rich (> 9%), MW 11–24 kD and forming mixed-coils [23]. RNA binding proteins (RBPs) were identified in *O. sinensis* and other fungi as RBPs are known to be involved in *M. robertsii*'s adaptation to cold conditions [24].

1.5 PCR verification of mating-type genes

To verify the presence/absence of the *MAT1-1-1* and *MAT1-2-1* genes in different isolates (Table S1), primers were designed based on the DNA sequences OCS_06642 for *MAT1-1-1* (MAT1-1F, 5′-TGGGTTTCTCACCACCTTGT-3′; and MAT1-1R, 5′-AGTTCATCACGGGCTCATTC-3′) and OCS_00196 for *MAT1-2-1* (MAT1-2F, 5′-CCACCGATCCAAGTCTCCT-3′; and MAT1-2R, 5′-CAGTTTCAGTCGCTGTCGTG-3′). Genomic DNA of each isolate was extracted from the liquid culture and used as a PCR template [25].

1.6 Phylogenetic analysis

To rebuild the evolutionary relationships of *O. sinensis*, 52 proteins broadly used for eukaryotic phylogenetic studies were selected [26]. Other ascomycetous fungal species with well-annotated genomes and clear sexuality patterns, i.e., either heterothallic or homothallic, were included in phylogenetic analysis. Putative orthologs from each fungus were identified as the highest full sequences with a minimum *E*-value of 1×10^{-20}. The basidiomycetous species *Ustilago maydis* was included to root the tree. The orthologous protein sequences were aligned with the program CLUSTAL X (ver 2.0) under default settings and the concatenated sequences were used for a Maximum-likelihood phylogenetic analysis using the software MEGA (ver. 5.0) with a Jones-Taylor-Thornton substitution model and 1000 bootstrap replications for phylogeny test [27].

1.7 Insect infection assay and fungal development induction

Late instar larvae of *Hepialus* spp. were reared individually in bottles at 18–20°C and fed with carrot slices. Spores of *O. sinensis* isolates Co18 and QH195-2 (Table S1) were harvested from one-month old liquid culture and washed twice with sterile distilled water. Cell walls and nuclei were stained with the fluorescent dyes Calcofluor white and 4′,6-diamidino-2-phenylindole (DAPI), respectively. The concentration of fungal cells was determined using a hemocytometer and adjusted to 1×10^{7} cells mL^{-1} with sterile water. Individual larva were anesthetized at –20°C for a few minutes and injected through the second proleg with 10 μL of fungal cell suspension. Insects injected with sterile water were used as controls. The treated larvae were returned to their bottles and at intervals were bled for microscopic examination of fungal development within the hemocoel. Following death, the stiff cadavers mummified by the fungus were buried in moisturized soil with their heads up to mimic natural conditions. Two batches (with 20 infected insects each) were incubated at different temperatures. One batch was incubated at 4°C for one month and then at 18°C for two months. The other batch of 20 cadavers was maintained at 18°C for three months. Field-collected parasitic complexes with immature fruiting bodies were continuously incubated in moisturized soil in cups at 20°C (light : dark control of 12 h : 12 h), to determine the possibilities of fruiting-body maturation and sexual structure development.

2 Results and discussion

2.1 General genome features

Using a combination of second generation methods, the genome of *O. sinensis* was sequenced to ~240-fold coverage which resulted in an estimation of 88.7% completeness of the genome. We found that due to repeat driven expansion (Figure 2; Table S2), the *O. sinensis* genome size is approximately three times larger (~120 Mb) than the median of other ascomycete insect pathogens. However, the number of *O. sinensis* curated protein coding genes is only 6972 as compared to more than 9500 genes in other fungi (Table 1). An expanded repeat-rich genome encoding fewer genes than relations is also a feature of some biotrophic obligate plant pathogens, e.g. *Blumeria graminis* [28] and *Puccinia graminis* [29], and the ectomycorrhizal fungus *Tuber melanosporum* [30].

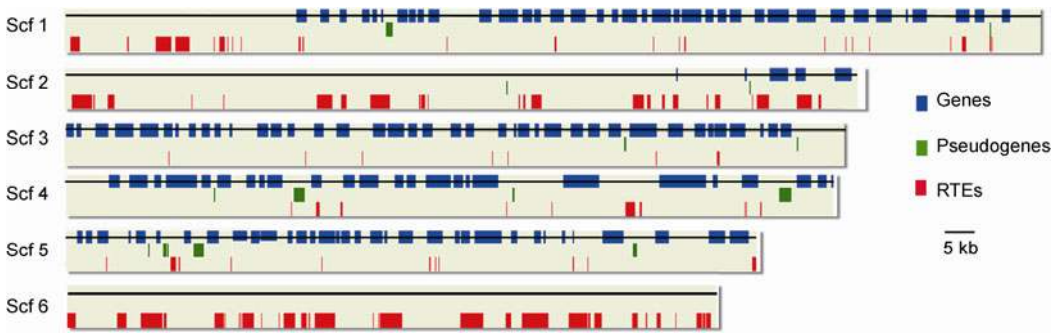

Figure 2 Distribution of genes, pseudogenes and retrotransposable elements (RTEs) on the six longest scaffolds of *O. sinensis*. Many scaffolds like Scf6 without any gene features are composed exclusively of retrotransposons whereas other scaffolds contain individual genes interspersed among retrotransposons.

We found that the genome defense process of repeat-induced-point-mutation is dysfunctional in *O. sinensis* (Figure S1), which has probably contributed to the massive proliferation of retrotransposable elements (RTEs) (Figure 2; Table S3), and thus genome size inflation. Unlike the obligate pathogen *B. graminis* [28], but similar to the ectomycorrhizal symbiont *T. melanosporum* [30], related RTEs were clustered together in gene-poor or gene-free regions of the *O. sinensis* genome indicative of repeated rounds of retrotransposition (Figure 2). As indicated by the number of retrotransposed and fragmented pseudogenes, retrotransposition could account for most of the observed gene losses in *O. sinensis* (Table S4). The categories of pseudogenized genes are consistent with a loss of capacity to adapt to heterogenous environments. In particular, the single *O. sinensis* nitrate reductase gene was pseudogenized, and the fungus also lacks nitric oxide reductase, suggesting it cannot assimilate nitrate. An inability to assimilate nitrate is also a feature of obligate plant pathogens [31].

2.2 Homothallism in the caterpillar fungus

We found that the *O. sinensis* genome has both *MAT1-1* (containing *MAT1-1-1*, OCS_06642; *MAT1-1-2*, OCS_06643 and *MAT1-1-3*, OCS_06644 genes) and *MAT1-2* (containing *MAT1-2-1*, OCS_00196) loci in separate scaffolds. PCR screening verified the presence of both *MAT1-1-1* and *MAT1-2-1* genes in 48 single ascospore isolates of six field-collected strains (Figure 3(a); Table S1). A staining assay verified that cells are uninucleate (Figure S2). Thus, *O. sinensis* is homothallic and can complete the sexual cycle in isolation. This is in contrast to the closely related insect pathogens *C. militaris* [17], *Beauveria* (*Cordyceps*) *bassiana* [18], and *Metarhizium* spp. (anamorphs of *Metacordyceps* spp.) [19] which are heterothallic. In the homothallic plant pathogen *Fusarium graminearum* the two Mat loci are fused together within the idiomorphic region [32], whereas in *A. nidulans* the Mat loci are in different chromosomes each with one flank of the idiomorphic region [33]. In contrast, the *MAT1-2* locus of *O. sinensis* is located in the conserved idiomorphic region, but the *MAT1-1* locus is isolated outside a MAT locus as a result of reshuffling by mobile genetic elements (Figure 3(b)). The only similar situation is found in *B. graminis*, where the *MAT1-2* locus is likewise not contained in a MAT region due to extensive retrotransposition [28]. Similar to its close relatives, the whole genome survey indicated that *O. sinensis* has all the genes currently known to be necessary for mating signaling, karyogamy, meiosis and fruiting-body development in the model fungi *A. nidulans* and *N. crassa* (Table S5).

A phylogenomic analysis confirmed that *O. sinensis* diverged from the heterothallic *Cordyceps* spp. (Figure 3(c)), suggesting that *O. sinensis* underwent an evolutionary transition from out-crossing to self-fertility that involved the acquisition of both mating types into the same haploid genome. The observation of further divergence into the heterothallic *Metarhizium* spp. after the origin of homothallic *O. sinensis* indicated that, like *Neurospora* spp. [34], multiple switches in mating systems have also occurred in the evolutionary history of insect pathogens (Figure 3(c)). It is likely that inbreeding is an adaptation by *O. sinensis* to a much more specialized lifestyle and extreme environmental conditions that produce a limited population size [34]. Consistent with this, *O. sinensis* has fewer (5 genes) heterokaryon incompatibility proteins than other insect pathogens (⩾15 genes) (Table S6) suggesting it encounters fewer genetically distinct individuals than its more opportunistic relatives and therefore does not need barriers to vegetative fusions [35].

2.3 Fruiting-body induction and maturation

Based on the discovery of *O. sinensis* homothallism, we tried to induce development of sexual fruiting-bodies in the

Table 1 Comparison of genome features between *O. sinensis* and other insect pathogens

Features	*O. sinensis*	*C. militaris*	*B. bassiana*	*M. robertsii*	*M. acridum*
Size (Mb)	~120	32.2	33.7	39	38.1
Coverage (fold)	241	147	76.6	100	107
G+C content (%)	46.1	51.4	51.5	51.5	50
Repeat rate (%)	37.98	3.04	2.03	0.98	1.52
Protein-coding genes	6972	9684	10366	10582	9849
Protein families (Protein No.)[a)]	2229 (4800)	2736 (6725)	3002 (7238)	2797 (7556)	2746 (6948)
Gene density (gene per Mb)	87	301	308	271	259
Exons per gene	2.6	3	2.7	2.8	2.7
Orphan proteins	742	885	723	363	284
Putative PHI genes[b)]	998	1547	2121	1828	1629
Pseudogenes	1110	102	304	363	440
NCBI accession	ANOV00000000	AEVU00000000	ADAH00000000	ADNJ00000000	ADNI00000000

a) InterProScan analysis data. b) Genes putatively involved in pathogen-host interactions (PHI).

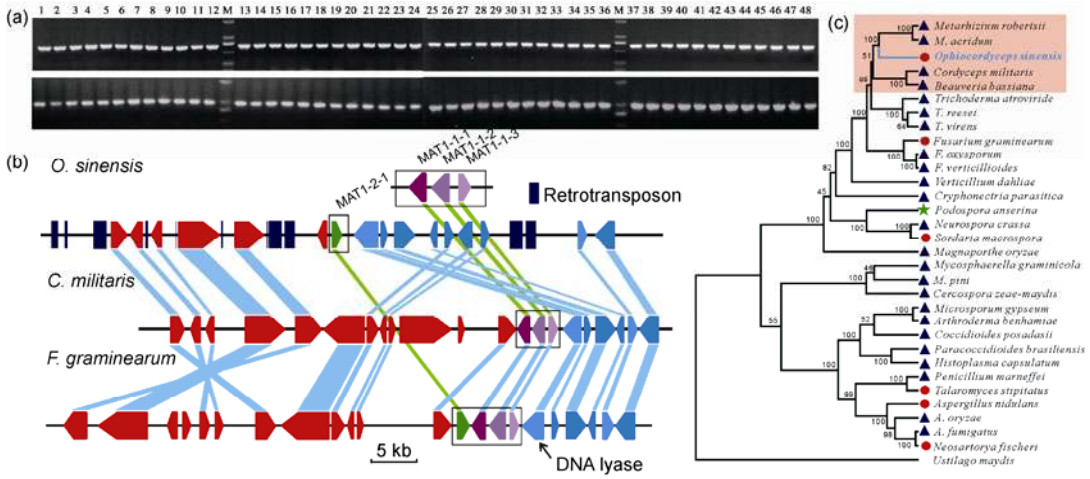

Figure 3 Genotyping and synteny analysis of *O. sinensis* MAT loci. (a) PCR examination showing the presence of both *MAT1-1-1* (top panel) and *MAT1-2-1* (bottom panel) mating type genes in 48 single ascospore isolates of *O. sinensis*. (b) Syntenic analysis of *O. sinensis*, *C. militaris* and *F. graminearum* idiomorphic regions. Mating-type loci are framed and the genes flanked at the same site of MAT locus are labeled in the same color. Multiple retrotransposons were observed in the *MAT1-2* locus of *O. sinensis*. (c) Evolutionary relationships among the heterothallic (labeled with purple triangles), homothallic (red circles) and pseudohomothallic (green stars) fungal species. Insect pathogens are highlighted in pink.

laboratory by injecting spores from different isolates into the body cavity (hemocoel) of late instar ghost moth caterpillars. Unlike other insect pathogens that grow rapidly *in insecta* and kill insects within 3–5 d of infection [13,36], there was a latent period of three to four weeks when the infected insects continued feeding and the hemocoel contained few fungal cells (Figure S3). Insects died about one month after injection followed by massive colonization ("mummification") of the cadaver by fungal cells. Incubating the mummified cadavers in soil for up to three months at different temperatures to mimic those in nature failed to induce the production of fruiting bodies (Figure S3). However, the field-collected specimens with newly initiated fruiting bodies successfully completed sexual development in the laboratory (Figure 4). Thus, the induction of sexual processes may be linked to the cryptic environmental factors specific to the Tibetan Plateau alpine ecosystem, but this is not required for sexual maturation.

2.4 Novel infection mechanisms

As described above, a remarkable characteristic of the repeat driven expansion of the *O. sinensis* genome is that a very similar process has occurred in some narrow host range obligate plant pathogens that have a biotrophic phase [28,29]. Also consistent with narrow lifestyle options, *O. sinensis* has > 2-fold fewer dehydrogenases (components of metabolic pathways) than other insect pathogens (103 vs. average 237) (Table S6). As with *B. graminis* [28], alcohol dehydrogenases and gluconate 5-dehydrogeases are particularly underrepresented in *O. sinensis* (Table S7). There has also been an overall loss of carbohydrate degrading enzymes with, for example, many fewer glycoside hydrolases (66 vs. average 139 in other insect fungi, and 199 in plant pathogens) [18]. Most of the missing enzymes are devoted to degradation of plant materials indicating that unlike related insect pathogens *O. sinensis* may be exclusively parasitic.

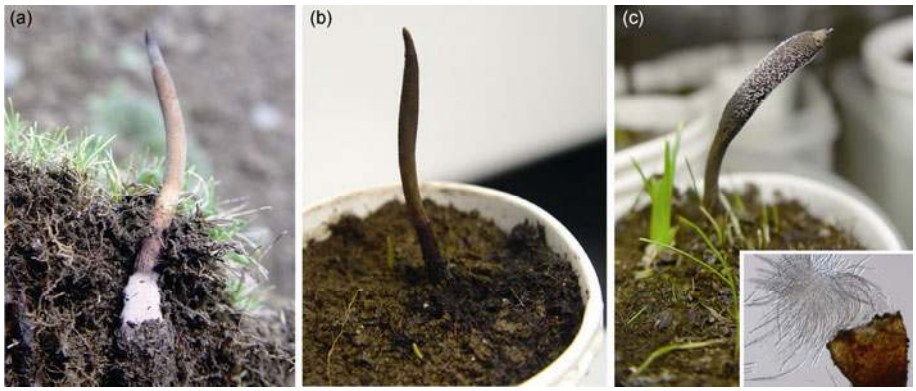

Figure 4 Maturation of *O. sinensis* sexual structures in the laboratory. A field-collected sample with an immature fruiting body (a) was incubated in soil (b) for up to two weeks, the fruiting body swelled (c) by producing mature sexual perithecia and asci (inset).

Typically, insect pathogenic fungi infect insects by breaching the cuticles using a combination of mechanical pressure [exerted by infection structures (appressoria and penetration pegs)] and cuticle degrading enzymes [37], which are greatly expanded in other insect pathogens [18]. However, *O. sinensis* gene families encoding cytochrome P450 (CYP) subfamily CYP52 enzymes (for metabolism of insect epicuticular lipids), cuticle degrading proteases (e.g. 2 trypsins in *O. sinensis* vs. average 21 in other insect pathogens; 17 subtilisins vs. 44 in other insect pathogens) and chitinases (9 vs. 19 in other insect pathogens) were greatly reduced in size. In addition, protein families involved in adhesion to cuticles and formation of appressoria were absent or reduced in *O. sinensis* (Table S6). These gene losses suggest that *O. sinensis* can no longer breach intact cuticle and instead probably infects insects orally or via the spiracles (breathing holes). Hydrolytic enzymes, particularly proteases, can elicit host immune defenses [38]. In which case, the reduced number of cuticle degrading enzymes in *O. sinensis* might be an adaptation to avoid detection by the host during the extended latent phase. Copy number reduction was also evident for genes encoding known pathogen-associated molecular patterns such as lectins (6 in *O. sinensis* vs. average 24 in other insect pathogens) (Table S6), further suggesting that selection for "stealth" (avoidance of host defenses) was a major force driving *O. sinensis* evolution [39]. *O. sinensis* encodes two orphan membrane-bound proteins (OCS_00468 and OCS_01831) with the characteristic collagenous G-X-Y repeat motif. *Metarhizium* covers itself with a collagen-like protein to evade the insect immune system [36].

2.5 Secondary metabolisms

Insect pathogens employ secondary metabolites for killing hosts [40]. Whereas biotrophic plant pathogens have lost all or most secondary metabolite genes [28], *O. sinensis* has multiple polyketide synthases, modular non-ribosomal peptide synthases and terpene cyclases for producing an array of secondary metabolites (Table S8). These most likely play roles after the latent period when the fungus is colonizing and killing the host, and they are also likely candidates for production of pharmacologically active compounds. *O. sinensis* encodes four terpenoid synthases (OCS_00133, OCS_00134, OCS_00704 and OCS_04474) and one terpenoid cyclase (OCS_00698) absent in other fungi, indicating that this "natural Viagra" produces novel bioactive terpenoids. However, many putative secondary metabolism clusters were conserved between *O. sinensis* and other insect pathogens (Figure S4), providing singular exceptions to the restructuring of the *O. sinensis* genome by repeat elements and suggesting that the physical linkage of secondary metabolite biosynthetic genes has strong adaptive significance for entomopathogenicity.

2.6 Cold adaptation

O. sinensis is the first psychrophilic fungus sequenced and we were curious about its strategies for temperature adaptation. *O. sinensis* cells contain large numbers of lipid droplets (Figure 5), as described in yeast species adapted to long-term survival in frigid conditions [41]. Compared to other fungi, *O. sinensis* has a similar range and number of genes involved in triacylglycerol and fatty acid biosynthesis but fewer lipases (22 vs. average 42 in other fungi) and fatty acid hydroxylases (6 vs. average 14 in other fungi) consistent with an emphasis on making rather than degrading lipids (Table S9). Like cold-adapted bacteria and plants [42], the *O. sinensis* genome was enriched in fatty acid desaturases (7 vs. average 3 in other fungi) (Table S9), suggesting that the fungus may respond to low temperatures by increasing membrane lipid unsaturation. *O. sinensis* encodes two orphan proteins (OCS_06184 and OCS_06507) structurally similar to insect antifreeze proteins (Table S9) that may be an adaptation to the dramatic freeze-thaw cycles of high-elevation environments. *O. sinensis* also has seven chloroperoxidases, whereas the average in other fungi is less than two (Table S6). Chloroperoxidases are able to catalyze oxygen transfers by sulfoxidation [43], and may therefore be an adaptation by *O. sinensis* to low oxygen levels at high altitudes.

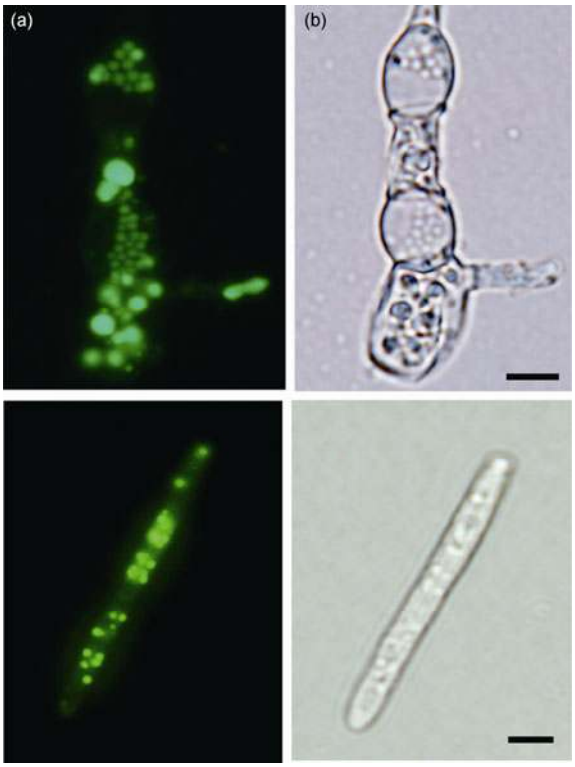

Figure 5 Lipid droplet staining of *O. sinensis* cells with the fluorescent dye Bodipy. (a) Hyphal cells harvested from a solid medium 70 d post inoculation. (b) A bar-shaped blastospore harvested from a 30-d old liquid culture. Right panels show bright-field images. Bar, 5 μm.

3 Conclusions

In conclusion, our data suggest that *O. sinensis* is a pathogen with a biphasic pathogenic mechanism beginning with stealth pathogenesis in early host instars and a lethal stage in late instars. It has been proposed that fungal lineages with large and flexible genomes are likely to adapt faster during co-evolution with hosts [44]. It is reasonable to assume that for the inbreeding *O. sinensis* the massive proliferation of RTEs provides a tradeoff between advantages of increased genetic variation independent of sexual recombination and deletion of genes dispensable for its specialized pathogenic lifestyle. Retrotransposition could therefore have a common role in the evolutionary origins of obligate host-specific fungi with a biphasic lifestyle. As *O. sinensis* has lost many genes for opportunism, future transitions away from its current lifestyle seem unlikely indicating that while retrotransposition may facilitate rapid adaptation, it may also contribute to stabile host interactions. Overall, the availability of this genome provides an unparalleled opportunity to develop a deeper understanding of how this unique pathogen interacts with insects within its ecosystem. It will also allow for more rapid identification of genes encoding biologically active molecules, and may lead to ways of sustainably utilizing this precious medical resource following further studies on the biotic and abiotic environmental factors that control sexual development.

This work was supported by the National Basic Research Program of China (2009CB118904), National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2007BAI32B00) and the National Natural Science Foundation of China (31225023). We thank the BGI at Shenzhen for sequencing service.

1 Zhang Y J, Li E W, Wang C S, et al. *Ophiocordyceps sinensis*, the flagship fungus of China: Terminology, life strategy and ecology. Mycology, 2012, 3: 2–10
2 Winkler D. Yartsa Gunbu (*Cordyceps sinensis*) and the fungal commodification of the rural economy in Tibet AR. Eco Bot, 2008, 62: 291–306
3 Dong C H, Yao Y J. Nutritional requirements of mycelial growth of *Cordyceps sinensis* in submerged culture. J Appl Microbiol, 2005, 99: 483–492
4 Paterson R R. *Cordyceps*: A traditional Chinese medicine and another fungal therapeutic biofactory? Phytochemistry, 2008, 69: 1469–1495
5 Mycology S R. Last stand for the body snatcher of the Himalayas? Science, 2008, 322: 1182
6 Holliday J, Cleaver M. Medicinal value of the caterpillar fungi species of the genus *Cordyceps* (Fr.) Link (Ascomycetes). A review. Int J Med Mushrooms, 2008, 10: 219–234
7 Stone R. Bhutan. Improbable partners aim to bring biotechnology to a Himalayan kingdom. Science, 2010, 327: 940–941
8 Cannon P F, Nigel L, Hywel-Jones, et al. Steps towards sustainable harvest of *Ophiocordyceps sinensis* in Bhutan. Biodivers Conserv, 2009, 18: 2263–2281
9 Li J F, Zou Z W, Liu X, et al. Biology of *Thitarodes pui* (Lepidoptera, Hepialidae) a host species of *Ophiocordyceps sinensis.* J Environ Entomol, 2011, 33: 195–202
10 Zheng P, Xia Y L, Zhang S W, et al. Genetics of *Cordyceps* and related fungi. Appl Microbiol Biotechnol, 2013, 97: 2797–2804
11 Ni M, Feretzaki M, Sun S, et al. Sex in fungi. Annu Rev Genet, 2011, 45: 405–430
12 Zhang S, Zhang Y J, Liu X Z, et al. Cloning and analysis of the MAT1-2-1 gene from the traditional Chinese medicinal fungus *Ophiocordyceps sinensis*. Fungal Biol, 2011, 115: 708–714
13 Duan Z B, Chen Y X, Huang W, et al. Linkage of autophagy to fungal development, lipid storage and virulence in *Metarhizium robertsii*. Autophagy, 2013, 9: 538–549
14 Li R, Zhu H, Ruan J, et al. *De novo* assembly of human genomes with massively parallel short read sequencing. Genome Res, 2010, 20: 265–272
15 Parra G, Bradnam K, Korf I. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics, 2007, 23: 1061–1067
16 Zhang Z, Cariero N, Zheng D, et al. PseudoPipe: An automated pseudogene identification pipeline. Bioinformatics, 2006, 22: 1437–1439
17 Zheng P, Xia Y L, Xiong C H, et al. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued Traditional Chinese Medicine. Genome Biol, 2011, 12: R116
18 Xiao G H, Ying S H, Zheng P, et al. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. Sci Rep, 2012, 2: 483
19 Gao Q, Jin K, Ying S H, et al. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. PLoS Genet, 2011, 7: e1001264
20 Dyer P S, O'Gorman C M. Sexual development and cryptic sexuality in fungi: Insights from *Aspergillus* species. FEMS Microbiol Rev, 2012, 36: 165–192
21 Medema M H, Blin K, Cimermancic P, et al. AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res, 2011, 39: W339–W346
22 Fedorova N D, Moktali V, Medema M H. Bioinformatics approaches and software for detection of secondary metabolic gene clusters. Methods Mol Biol, 2012, 944: 23–45
23 Venketesh S, Dayananda C. Properties, potentials, and prospects of antifreeze proteins. Crit Rev Biotechnol, 2008, 28: 57–82
24 Fang W G, St. Leger R J. RNA binding proteins mediate the ability of a fungus to adapt to the cold. Environ Microbiol, 2010, 12: 810–820
25 Zhang Y J, Xu L L, Zhang S, et al. Genetic diversity of *Ophiocordyceps sinensis*, a medicinal fungus endemic to the Tibetan Plateau: Implications for its evolution and conservation. BMC Evol Biol, 9: 290
26 Floudas D, Binder M, Riley R, et al. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science, 2012, 336: 1715–1719
27 Tamura K, Peterson D, Peterson N, et al. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol, 2011, 28: 2731–2739
28 Spanu P D, Abbott J C, Amselem J, et al. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science, 2010, 330: 1543–1546
29 Duplessis S, Cuomo C A, Lin Y C, et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci USA, 2011, 108: 9166–9171
30 Martin F, Kohlrt A, Murat C, et al. Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature, 2010, 464: 1033–1038
31 Spanu P D. The genomics of obligate (and nonobligate) biotrophs. Annu Rev Phytopathol, 2012, 50: 91–109
32 Ma L J, van de Does H C, Borkovivh K A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature, 2010, 464: 363–373
33 Galagan J E, Calvo S E, Cuomo C, et al. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*.

Nature, 2005, 438: 1105–1115
34 Nygren K, Strandberg R, Wallberg A, et al. A comprehensive phylogeny of *Neurospora* reveals a link between reproductive mode and molecular evolution in fungi. Mol Phylogenet Evol, 2011, 59: 649–663
35 Glass N L, Dementhon K. Non-self recognition and programmed cell death in filamentous fungi. Curr Opin Microbiol, 2006, 9: 553–558
36 Wang C S, St. Leger R J. A collagenous protective coat enables *Metarhizium anisopliae* to evade insect immune responses. Proc Natl Acad Sci USA, 2006, 103: 6647–6652
37 St. Leger R J, Wang C S. Genetic engineering of fungal biocontrol agents to achieve greater efficacy against insect pests. Appl Microbiol Biotechnol, 2010, 85: 901–907
38 St Leger R J, Joshi L, Bidochka M J, et al. Construction of an improved mycoinsecticide overexpressing a toxic protease. Proc Natl Acad Sci USA, 1996, 93: 6349–6354
39 Rappleye C A, Goldman W E. Fungal stealth technology. Trends Immunol, 2008, 29: 18–24
40 Wang B, Kang Q J, Lu Y Z, et al. Unveiling the biosynthetic puzzle of destruxins in *Metarhizium* species. Proc Natl Acad Sci USA, 2012, 109: 1287–1292
41 Amaretti A, Raimondi S, Sala M, et al. Single cell oils of the cold-adapted oleaginous yeast *Rhodotorula glacialis* DBVPG 4785. Microb Cell Fact, 2010, 9: 73
42 D'Amico S, Collins T, Marx J C, et al. Psychrophilic microorganisms: Challenges for life. EMBO Rep, 2006, 7: 385–389
43 Hofrichter M, Ullrich R, Pecyna M J, et al. New and classic families of secreted fungal heme peroxidases. Appl Microbiol Biotechnol, 2010, 87: 871–897
44 Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: Why bigger can be better. Nat Rev Microbiol, 2012, 10: 417–430

Supporting Information

Figure S1 Analysis of repeat-induced point mutation (RIP) in *O. sinensis*.
Figure S2 Examination of the number of n\uclei in *O. sinensis* cells harvested from liquid culture.
Figure S3 Failure to induce *O. sinensis* fruiting-body production in the laboratory.
Figure S4 Syntenic relationship analysis of gene clusters involved in the biosynthesis of different secondary metabolites between *O. sinensis* and other insect pathogens.
Table S1 The single ascospore isolates used for mating type characterization.
Table S2 Repeat rate analysis of *O. sinensis* genome-sequencing data.
Table S3 Repeat-masking analysis of the assembled *O. sinensis* genome.
Table S4 Comparison of pseudogene numbers in *O. sinensis* and other insect pathogens.
Table S5 Sexuality-related genes functionally verified in *A. nidulans* and *N. crassa*, and their putative orthologs in *O. sinensis* and other insect pathogens.
Table S6 Comparison of selected protein families between *O. sinensis* and other insect pathogens.
Table S7 Comparison of dehydrogenases encoded in the insect pathogens and plant obligate pathogen *B. graminis*.
Table S8 The core proteins involved in the biosynthesis of secondary metabolites in *O. sinensis* and other insect pathogens.
Table S9 Comparison of proteins putatively involved in cold responses in *O. sinensis* and other insect pathogens.

The supporting information is available online at csb.scichina.com and www.springerlink.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.