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 anti-freeze 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].

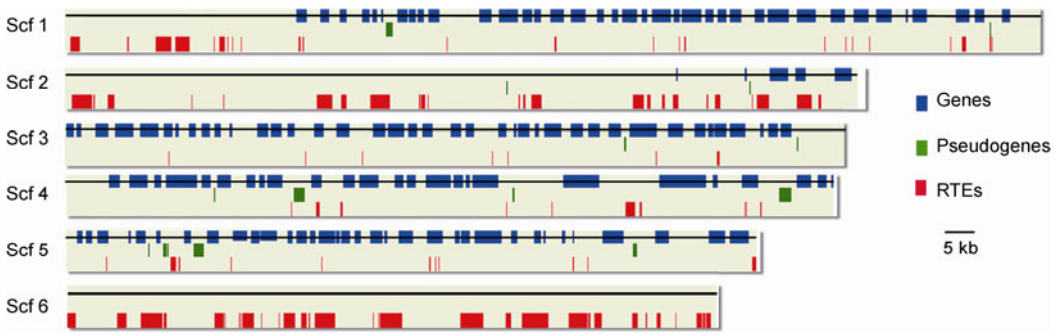

Figure 2 Distribution of genes, pseudogenes and retrotransposable elements (RTEs) on the six longest scaffolds of *O. sinensis*. Many scaffolds like Scf6 without any gene features are composed exclusively of retrotransposons whereas other scaffolds contain individual genes interspersed among retrotransposons.

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).

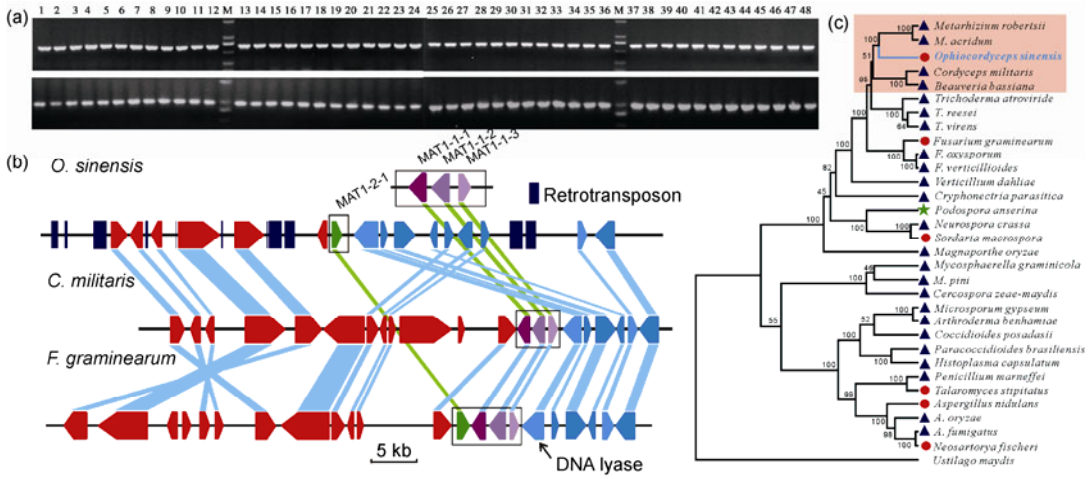

Figure 3 Genotyping and synteny analysis of *O. sinensis* MAT loci. (a) PCR examination showing the presence of both *MAT1-1-1* (top panel) and *MAT1-2-1* (bottom panel) mating type genes in 48 single ascospore isolates of *O. sinensis*. (b) Syntenic analysis of *O. sinensis*, *C. militaris* and *F. 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.

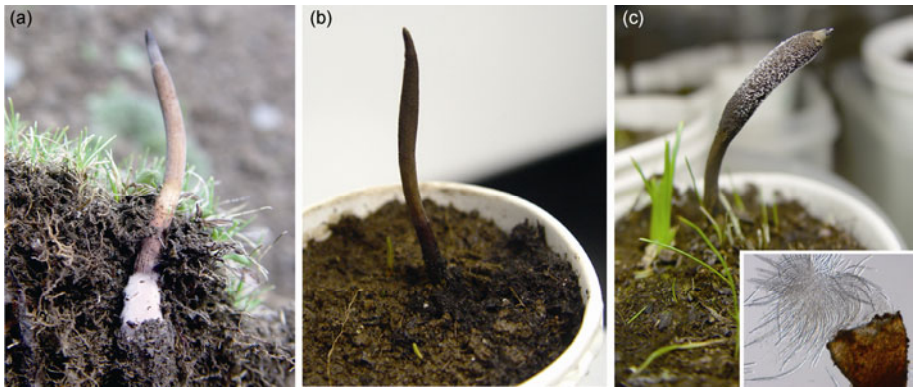

Figure 4 Maturation of *O. sinensis* sexual structures in the laboratory. A field-collected sample with an immature fruiting body (a) was incubated in soil (b) for up to two weeks, the fruiting body swelled (c) by producing mature sexual perithecia and asci (inset).

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.

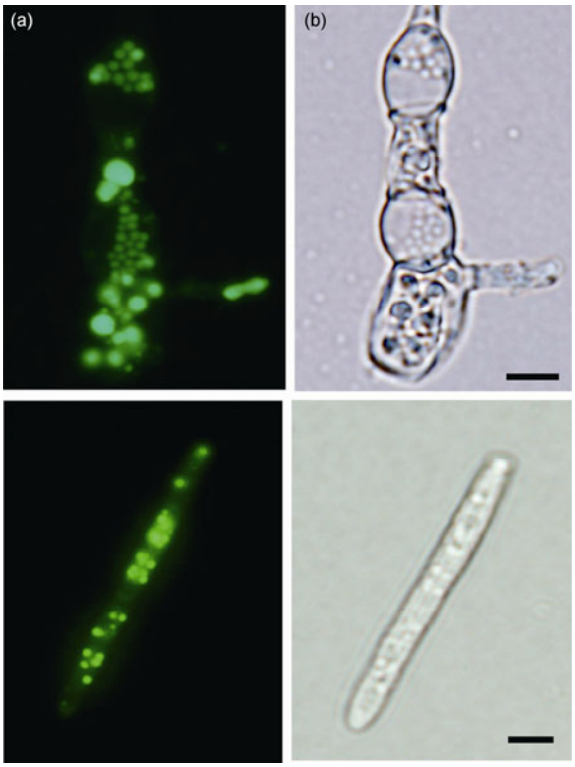

Figure 5 Lipid droplet staining of *O. sinensis* cells with the fluorescent dye Bodipy. (a) Hyphal cells harvested from a solid medium 70 d post inoculation. (b) A bar-shaped blastospore harvested from a 30-d old liquid culture. Right panels show bright-field images. Bar, 5 μm.

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.

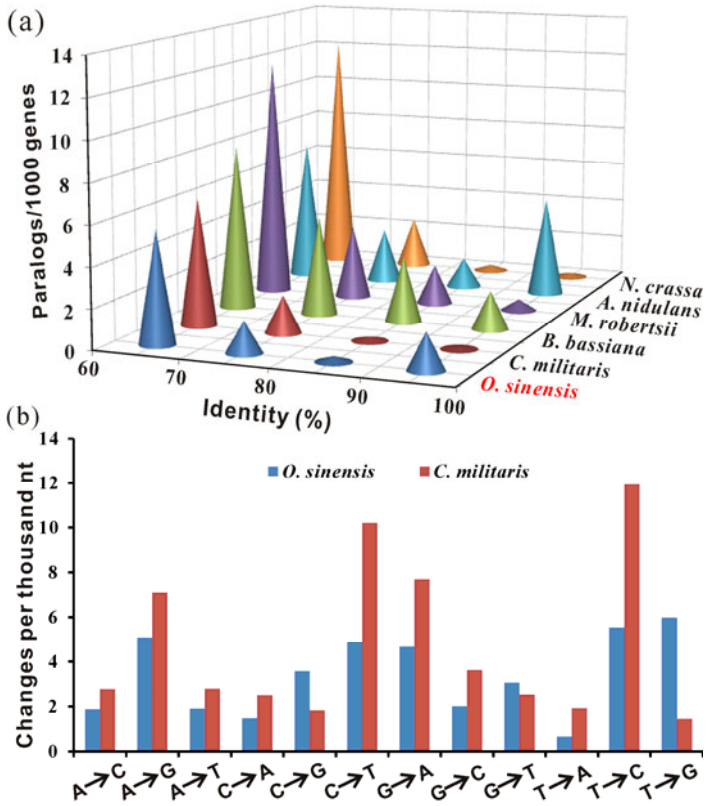

Figure S1 Analysis of repeat-induced point mutation (RIP) in *O. sinensis*. RIP is a fungal genome defense mechanism guarding against transposon invasion. RIP mutates the sequence of repeated DNA introducing a C:G to T:A mutation bias. (a) Distribution of paralogous gene numbers with different level of nucleotide similarity in *O. sinensis* and other fungi. We found 12 pairs with > 90% nucleotide sequence similarities in *O. sinensis* in contrast to only one pair each in RIP-effective species *Neurospora crassa* and *Cordyceps militaris*. (b) Analysis and comparison of nucleotide mutation rates of the paired *O. sinensis* and *C. militaris* genes showing >80% nucleotide identities. In contrast to *C. militaris*, *O. sinensis* lacks the C:G to T:A mutation bias characteristic of RIP.

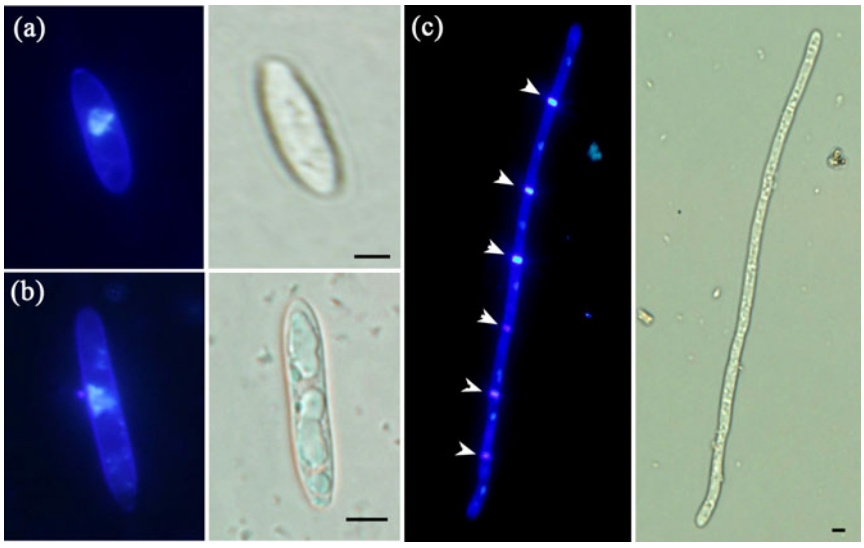

Figure S2 Examination of the number of nuclei in *O. sinensis* cells harvested from liquid culture. The cells were double stained with the Calcoflour White and 4',6-diamidino-2-phenylindole (DAPI). Panels (a), (b) and (c) show a yeast-like blastospore, a bar-shaped blastospore and mycelium, respectively. Arrows in Panel (c) point to a hyphal septum. The right panels show bright field images. Bar, 5 μm.

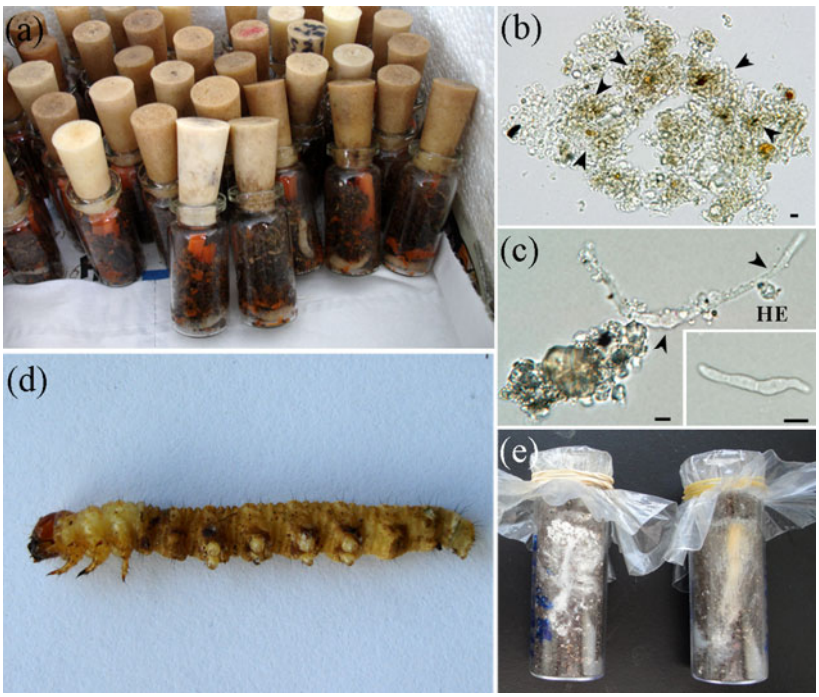

Figure S3 Failure to induce *O. sinensis* fruiting-body production in the laboratory. (a) Late instar larvae of ghost moth caterpillars were injected with 10 μl of fungal cells (1×10^7 cells per ml of water). (b) Spores were encapsulated and melanized (arrowed) by insect hemocytes 24 hours post injection. (c) Three weeks post infection, fungal cells (arrowed) were found growing within the treated insect hemocoel but not in insects from the control experiment. Inset: fungal cells observed in a newly killed insect one month post injection. Bar, 5 μm. (d) A dead insect colonized by *O. sinensis*. (e) Insect cadavers incubated in soil for up to three months failed to produce fruiting bodies. Left sample first treated at 4 °C for one month and then two months at 18 °C; Right sample treated for three months at 18 °C.

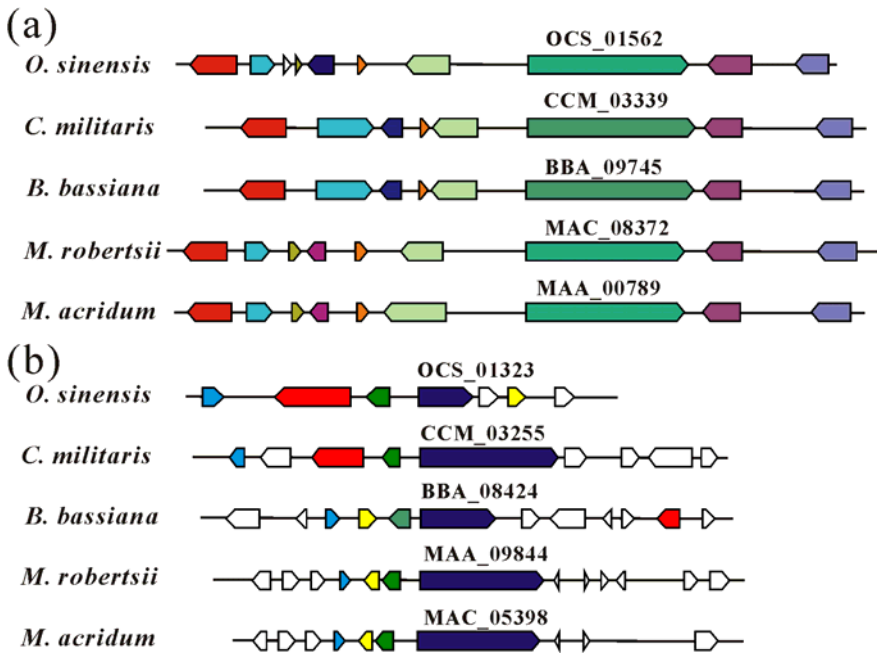

Figure S4 Syntenic relationship analysis of gene clusters involved in the biosynthesis of different secondary metabolites between *O. sinensis* and other insect pathogens. (a) A polyketide synthase gene (OCS_01562) cluster, and (b) a non-ribosomal peptide synthetase gene (OCS_01323) cluster is conserved in four insect pathogens. In each panel, orthologous genes are labeled in the same color. The genes labeled in white are orphans.

Table S1 The single ascospore isolates used for mating type characterization

Ser. No.	Isolate No.	Collection source
1	XZ49-1	Changdu, Tibet
2	XZ49-2	Changdu, Tibet
3	XZ49-3	Changdu, Tibet
4	XZ49-4	Changdu, Tibet
5	XZ49-5	Changdu, Tibet
6	XZ49-6	Changdu, Tibet
7	XZ49-7	Changdu, Tibet
8	XZ49-8	Changdu, Tibet
9	XZ49-9	Changdu, Tibet
10	XZ49-10	Changdu, Tibet
11	XZ49-11	Changdu, Tibet
12	XZ49-12	Changdu, Tibet
13	XZ49-13	Changdu, Tibet
14	XZ51-1	Changdu, Tibet
15	XZ51-2	Changdu, Tibet
16	XZ51-3	Changdu, Tibet
17	XZ51-4	Changdu, Tibet
18	XZ51-5	Changdu, Tibet
19	XZ51-6	Changdu, Tibet
20	QH195-2	Yushu, Qinhai
21	XZ36-3	Changdu, Tibet
22	XZ36-4	Changdu, Tibet
23	XZ36-5	Changdu, Tibet
24	XZ36-6	Changdu, Tibet
25	XZ157-1	Naqu, Tibet
26	XZ157-2	Naqu, Tibet
27	XZ157-3	Naqu, Tibet
28	XZ157-4	Naqu, Tibet
29	XZ157-5	Naqu, Tibet
30	XZ157-6	Naqu, Tibet
31	XZ157-7	Naqu, Tibet
32	XZ157-8	Naqu, Tibet
33	XZ157-9	Naqu, Tibet
34	XZ157-10	Naqu, Tibet
35	XZ157-11	Naqu, Tibet
36	XZ157-12	Naqu, Tibet
37	XZ157-13	Naqu, Tibet
38	XZ157-14	Naqu, Tibet
39	XZ157-15	Naqu, Tibet
40	XZ157-16	Naqu, Tibet
41	XZ141-1	Naqu, Tibet
42	XZ141-2	Naqu, Tibet
43	XZ141-3	Naqu, Tibet
44	XZ141-4	Naqu, Tibet
45	XZ141-5	Naqu, Tibet
46	XZ141-6	Naqu, Tibet
47	XZ141-7	Naqu, Tibet
48	XZ141-8	Naqu, Tibet

Table S2 Repeat rate analysis of *O. sinensis* genome-sequencing data

	Number	**Length**	**Repeat number**	**Repeat length**	**No. ratio (%)**	**Length ratio (%)**
Short contigs	3,346,441	93,452,497	1,565,042	43,214,452	46.76	46.24
Long contigs	48,133	20,840,395	651	202,264	1.35	0.97
Total	3,394,574	114,292,892	1,565,693	43,416,716	46.12	37.98

Table S3 Repeat-masking analysis of the assembled *O. sinensis* genome

Class	**No.**	**Ratio (% of total No.)**
DNA/En-Spm	1	0.0030
DNA/Harbinger	33	0.0983
DNA/hAT-Tip100	2	0.0060
DNA/Helitron	2	0.0060
DNA/IS4EU	4	0.0119
DNA/MuDR	276	0.8223
DNA/P	12	0.0358
DNA/TcMar-Fot1	147	0.4380
DNA/TcMar-Marin	35	0.1043
DNA/TcMar-Pogo	28	0.0834
DNA/TcMar-Tc1	28	0.0834
DNA/TcMar-Tc2	8	0.0238
DNA/TcMar-Tigge	1	0.0030
LINE/CR1	36	0.1073
LINE/I	1	0.0030
LINE/Jockey	47	0.1400
LINE/L1	7	0.0209
LINE/L2	43	0.1281
LINE/LOA	14	0.0417
LINE/Penelope	9	0.0268
LINE/R1	19	0.0566
LINE/R2	5	0.0149
LINE/Rex-Babar	1	0.0030
LINE/RTE-BovB	9	0.0268
LINE/Tad1	1365	4.0667
LTR/Copia	8353	**24.8860**
LTR/Delta	1	0.0030
LTR/ERV1	50	0.1490
LTR/ERVK	9	0.0268
LTR/ERVL	1	0.0030
LTR/Foamy	6	0.0179
LTR/Gypsy	15357	**45.7530**
LTR/Lenti	1	0.0030
LTR/Pao	8	0.0238
RC/Helitron	239	0.7121
Tandem repeats	7407	**22.0676**

Table S4 Comparison of pseudogene numbers in *O. sinensis* and other insect pathogens

Pseudogene type	*O. sinensis*	*C. militaris*	*B. bassiana*	*M. robertsii*	*M. acridum*
DUP	143	9	72	129	254
PSSD	111	0	62	122	104
FRAG	856	139	170	112	82
Total	**1,110**	**148**	**304**	**363**	**440**

Note: DUP, duplicated pseudogenes; PSSD, processed (retrotransposed) pseudogenes; FRAG, fragmented pseudogenes.

Table S5 Sexuality-related genes functionally verified in *A. nidulans* and *N. crassa*, and their putative orthologs in *O. sinensis* and other insect pathogens

Function	*A. nidulans*	*N. crassa*	*O. sinensis*	*C. militaris*	*B. bassiana*	*M. robertsii*	*M. acridum*
Mating process							
Mating-type MAT1-1-1 (MatB)	AN2755	NCU01958	OCS_06642	CCM_06523	BBA_07733	MAA_03718	
Mating-type MAT1-1-2			OCS_06643	CCM_09679		MAA_10977	
Mating-type MAT1-1-3			OCS_06644			MAA_03719	
Mating-type MAT1-2-1 (MatA)	AN4734	NCU01960	OCS_00196				MAC_05350
Carboxypeptidase alpha-factor processing	AN1384	NCU04316	OCS_02355	CCM_02046	BBA_04810	MAA_02969	MAC_03847
Endoprotease for alpha-factor processing	AN3583	NCU03219	OCS_02254	CCM_04990	BBA_09698	MAA_05263	MAC_02797
Dipeptidyl aminopeptidase alpha-factor processing	AN2946	NCU02515	OCS_03713	CCM_07274	BBA_07034	MAA_03364	MAC_07820
Dipeptidyl aminopeptidase for a-factor processing	AN8044	NCU00481	OCS_02388	CCM_02661	BBA_04103	MAA_05780	MAC_03112
CAAX prenyl protease a-factor C-terminal processing	AN6528	NCU11314	OCS_03088	CCM_05829	BBA_08336		MAC_01887
CAAX-farnesyltransferase beta subunit; a-factor modification	AN2002	NCU05999	OCS_03113	CCM_05066	BBA_03884	MAA_02158	MAC_02644
CAAX-farnesyltransferase alpha subunit	AN6162	NCU00034	OCS_06299	CCM_05030	BBA_01915	MAA_07087	MAC_07177
ATP-dependent efflux pump for a-factor like pheromone	AN2300	NCU07546	OCS_03625	CCM_00623	BBA_08160	MAA_05864	MAC_03697
Mating signaling							
Alpha-subunit G protein	AN0651	NCU06493	OCS_06305	CCM_07236	BBA_01892	MAA_03488	MAC_04984
Beta-subunit G protein	AN0081	NCU00440	OCS_04898	CCM_09461	BBA_03738	MAA_02470	MAC_01638
Gamma-subunit G protein	AN2742	NCU00041	OCS_06260	CCM_01794	BBA_05932	MAA_05991	MAC_08659
Mitogen-activated protein kinase	AN3719	NCU02393	OCS_04717	CCM_01235	BBA_01244	MAA_04503	MAC_00098
Pheromone adaptation feedback response	AN7252	NCU00455	OCS_06899	CCM_09426	BBA_03777	MAA_02467	MAC_01635
Pheromone Receptor (for a-factor like pheromone)	AN7743	NCU00138	OCS_06804	CCM_01499	BBA_04271	MAA_05941	MAC_00610
Pheromone Receptor (for alpha-factor like pheromone)	AN2520	NCU05758	OCS_05165	CCM_06752	BBA_04483	MAA_00341	MAC_02467
Serine/threonine protein kinase MKK	AN3422	NCU04612	OCS_00626	CCM_05558	BBA_04254	MAA_04421	MAC_02314
Serine/threonine protein kinase MKKK	AN2269	NCU06182	OCS_04114	CCM_07511	BBA_02280	MAA_03529	MAC_05029
Serine/threonine protein kinase MKKKK	AN2067	NCU03894	OCS_01179	CCM_05107	BBA_07438	MAA_02118	MAC_02685
Transcriptional Activator, homeodomain DNA binding	AN2290	NCU00340	OCS_04523	CCM_01444	BBA_01078	MAA_05929	MAC_00598
Karyogamy and meiosis							
1,3-beta-glucan synthase subunit, Sp Mei4 target gene	AN3729	NCU06871	OCS_05817	CCM_01185	BBA_10207	MAA_04456	MAC_01688
ABC transporter required fo a-factor transport	AN2300	NCU07546	OCS_03625	CCM_00623	BBA_08160	MAA_05864	MAC_03697
Amino acid permease involved in sexual differentiation	AN5678	NCU05830	OCS_02119	CCM_08184	BBA_00594	MAA_04108	MAC_08710
APC component	AN0905	NCU00494	OCS_04067	CCM_07395	BBA_08125	MAA_03573	MAC_08351
APC component	AN2772	NCU05901	OCS_05008	CCM_01001	BBA_01542	MAA_04652	MAC_09483
APC component	AN4735	NCU01963	OCS_00199	CCM_06526	BBA_07736	MAA_03722	MAC_05347
APC component	AN8013	NCU01174	OCS_02500	CCM_03205	BBA_02533	MAA_09174	MAC_05975
APC component, bimA	AN6138	NCU00217	OCS_06191	CCM_07521	BBA_02267	MAA_03443	MAC_02516
APC regulator	AN0814	NCU02616	OCS_02978	CCM_05397	BBA_01439	MAA_00735	MAC_06221
APC regulator	AN2965	NCU01269	OCS_04570	CCM_03958	BBA_07253	MAA_03661	MAC_06443
ARN family of transporters for siderophore-iron chelates	AN5378		OCS_04427	CCM_08779	BBA_07619	MAA_06654	MAC_09677

(*To be continued on the next page*)

(*Continued*)

Function	*A. nidulans*	*N. crassa*	*O. sinensis*	*C. militaris*	*B. bassiana*	*M. robertsii*	*M. acridum*
bZip TF, involved in regulation of meiosis	AN6849		OCS_05442	CCM_09124	BBA_04813	MAA_02972	MAC_08931
Calmodulin-dependent protein kinase	AN3065	NCU02283	OCS_03191	CCM_01313	BBA_00508	MAA_04599	MAC_04622
Chitin synthase, pheromone inducible	AN4566	NCU04251	OCS_04853	CCM_08511	BBA_03590	MAA_03168	MAC_08014
Chromosome segregation, kinetochore-associated Ndc80	AN4969	NCU03899	OCS_03932	CCM_00302	BBA_03433	MAA_02015	MAC_03415
Cohesin	AN6364	NCU07554	OCS_01053	CCM_01887	BBA_01045	MAA_05900	MAC_03734
Phosphocholine cytidylyltransferase	AN1357	NCU03880	OCS_00305	CCM_08417	BBA_06142	MAA_00035	MAC_07726
C-type cyclin	AN2172	NCU01563	OCS_02857	CCM_01137	BBA_04475	MAA_05354	MAC_01841
Cyclin-like protein interacts with Sc Pho85	AN4984	NCU07293	OCS_03462	CCM_07026	BBA_04435	MAA_03884	MAC_04184
Cytochrome-c oxidoreductase	AN3901	NCU08272	OCS_05124	CCM_02991	BBA_06880	MAA_01131	MAC_06772
DNA-binding helix-hairpin-helix protein	AN1237	NCU02741	OCS_03384	CCM_06822	BBA_02058	MAA_04013	MAC_04218
Dynactin complex, homologue of Sc NIP100	AN6323	NCU03483	OCS_04663	CCM_04587	BBA_05442	MAA_08863	MAC_07228
ER associated glutathione S-transferase	AN0629	NCU05706	OCS_02316	CCM_01615	BBA_03252	MAA_08951	MAC_06695
Fork head domain TF, meiotic regulator	AN8858	NCU06173	OCS_06910	CCM_02646	BBA_04116	MAA_05810	MAC_03141
Formin	AN6523	NCU01431	OCS_02615	CCM_07806	BBA_06491	MAA_01475	MAC_03937
GEF involved in conjugation; related to Sc CDC25	AN2130	NCU06500	OCS_06187	CCM_05772	BBA_06377	MAA_03615	MAC_06489
Glucosamine acetyl transferase involved in cell cycle progression	AN8706	NCU01902	OCS_05808	CCM_06566	BBA_02625	MAA_08730	MAC_05880
GTP binding (alpha-1 subunit) involved in conjugation	AN3090	NCU06729	OCS_06648	CCM_00555	BBA_06035	MAA_05603	MAC_01131
Hect domain E3 ubiquitin-protein ligase	AN0444	NCU06756	OCS_03053	CCM_04538	BBA_05494	MAA_07661	MAC_06530
HMG-box TF, target of pheromone signaling	AN3667	NCU09387	OCS_06292	CCM_06680	BBA_10247	MAA_00248	MAC_04576
Holliday junction resolvase	AN3118	NCU07457	OCS_04743	CCM_06887	BBA_02088	MAA_03931	MAC_05298
Inducer of meiosis, S/T kinase	AN6243	NCU01498	OCS_02952	CCM_02359	BBA_00996	MAA_05403	MAC_01790
involved in Meiosis II nuclear division, Sp Mei4 target gene	AN3931	NCU02540	OCS_04125	CCM_06283	BBA_09711	MAA_04211	MAC_03800
kinesin-like motor required for karyogamy	AN6340	NCU04581	OCS_02056	CCM_05847	BBA_08352	MAA_04756	MAC_02000
Long chain fatty acid elongation enzyme	AN8117	NCU06694	OCS_02273	CCM_03752	BBA_04877	MAA_07869	MAC_05925
MADS-box domain TF, pheromone receptor activator	AN8676	NCU07430	OCS_04793	CCM_03292	BBA_06763	MAA_07379	MAC_00861
Mating-type M-specific polypeptide Mc, HMG-box TF	AN1962	NCU03481	OCS_05413	CCM_04586	BBA_05443	MAA_10631	MAC_07229
Meiosis-specific transcriptional activator	AN6015	NCU09915	OCS_01689	CCM_04034	BBA_05141	MAA_06439	MAC_08428
Meiotic B-type cyclin	AN2137	NCU01242	OCS_06428	CCM_05610	BBA_00832	MAA_07313	MAC_00492
M-factor farnesyl cysteine carboxyl methyltransferase	AN6162	NCU00034	OCS_06300	CCM_05030	BBA_01915	MAA_07087	MAC_07177
Microtubule-binding protein	AN2862	NCU00243	OCS_00496	CCM_02770	BBA_04004	MAA_09680	MAC_09803
Phosphoglucomutase	AN2867	NCU10058	OCS_02582	CCM_02278	BBA_04232	MAA_04308	MAC_03213
Polo kinase	AN1560	NCU09258	OCS_06003	CCM_06972	BBA_08042	MAA_00136	MAC_07512
P-type ATPase sodium pump	AN6642	NCU05046	OCS_04959	CCM_04844	BBA_09050	MAA_08450	MAC_06648
Ras	AN0182	NCU08823	OCS_05591	CCM_06996	BBA_04387	MAA_06380	MAC_01355
Required for autophagic vesicle delivery to vacuole	AN5131	NCU01545	OCS_05823	CCM_07484	BBA_02302	MAA_02674	MAC_08764
Meiosis-specific topoisomerase Spo11	AN8259	NCU01120	OCS_03978	CCM_09527			
Rho-GEF	AN4719	NCU00668	OCS_06478	CCM_03356	BBA_07881	MAA_07421	MAC_00815
Riboflavin synthase, alpha subunit	AN4231	NCU07456	OCS_04744	CCM_06886	BBA_02087	MAA_03932	MAC_05297
Ribose-5-phosphate isomerase	AN2440	NCU07608	OCS_06483	CCM_00462	BBA_03239	MAA_02753	MAC_08154
RING-type ubiquitin ligase, FYVE finger domain	AN0627	NCU08974	OCS_02185	CCM_05861	BBA_08364	MAA_05768	MAC_03100
RNA binding protein required for meiotic recombination	AN9090	NCU00768	OCS_05666	CCM_03581	BBA_03049	MAA_07476	MAC_07476
RNA-binding protein involved in meiosis	AN6494	NCU00118	OCS_05556	CCM_05815	BBA_09975	MAA_03639	MAC_06466
RNA-binding protein involved in meiosis, Sp Mei4 target	AN7700	NCU00556	OCS_00350	CCM_00789	BBA_02405	MAA_03790	MAC_00971
SAM domain, similar to Sc STE50	AN7252	NCU00455	OCS_06899	CCM_09426	BBA_03777	MAA_02467	MAC_01635
Septin	AN1394	NCU03515	OCS_00123	CCM_00216	BBA_08598	MAA_02556	MAC_03632
Serine carboxypeptidase, degrades extracellular P-factor	AN2555	NCU06720	OCS_02487	CCM_01903	BBA_00850	MAA_09950	MAC_02763
Serine/threonine protein kinase, negative regulator of meiosis	AN4935	NCU04990	OCS_04406	CCM_00683	BBA_08648	MAA_09884	MAC_02282

(*To be continued on the next page*)

(*Continued*)

Function	*A. nidulans*	*N. crassa*	*O. sinensis*	*C. militaris*	*B. bassiana*	*M. robertsii*	*M. acridum*
SH3-domain protein	AN2995	NCU00696	OCS_00737	CCM_00772	BBA_02425	MAA_03765	MAC_07579
Spindle pole body component	AN5521	NCU04535	OCS_01011	CCM_00173	BBA_08305	MAA_05499	MAC_07346
Spindle pole body component	AN5618	NCU09871	OCS_00102	CCM_04480	BBA_02843	MAA_02280	MAC_03427
Splicesome component	AN1861	NCU00361	OCS_01066	CCM_01429	BBA_01066	MAA_05886	MAC_03719
Sporulation regulated septin	AN4667	NCU02464	OCS_02832	CCM_05213	BBA_05547	MAA_10305	MAC_04367
Sporulation regulated septin	AN8182	NCU03795	OCS_06471	CCM_09142	BBA_04748	MAA_07332	MAC_00472
Sporulation-specific enzyme required for spore wall maturation	AN2705		OCS_04544	CCM_01728	BBA_09587	MAA_09988	MAC_08553
ubiquitin-conjugating enzyme, chromosome segregation	AN5495	NCU11509	OCS_01386	CCM_05530	BBA_00216	MAA_00522	MAC_07614
Vacuolar ATPase Ca2+ pump	AN1189	NCU04736	OCS_01528	CCM_04692	BBA_06258	MAA_00012	MAC_02382
Xyulose kinase	AN8790	NCU11353	OCS_05208	CCM_09684	BBA_02252	MAA_03458	MAC_02531
Zn finger transcription factor, unknown function	AN3447	NCU05007	OCS_05234	CCM_08238	BBA_00653	MAA_06676	MAC_07713
Fruiting body development							
Binuclear zinc transcription factor	AN5170	NCU07392	OCS_05114	CCM_00123	BBA_01641	MAA_05237	MAC_05412
Cys6-Zn2 transcriptional activator	AN1848	NCU07392	OCS_02179	CCM_02196	BBA_04319	MAA_04374	MAC_06154
GATA-transcription factor	AN3152	NCU01154	OCS_03409	CCM_00560	BBA_06064	MAA_05601	MAC_01133
Modulator of pheromone-inducible gene expression	AN8676	NCU07430	OCS_04793	CCM_03292	BBA_06763	MAA_07379	MAC_00861
Mutanase, cell wall turnover during sexual development	AN7349						
Regulator of myosin function	AN0135	NCU06821	OCS_05043	CCM_02480	BBA_09803	MAA_02633	MAC_03554
Repressor of b mating type regulated genes	AN8211	NCU01238	OCS_04623	CCM_07966	BBA_06597	MAA_01434	MAC_04053
Signalosome subunit 4, regulation of sexual development	AN1539	NCU07361	OCS_00445	CCM_08174	BBA_00583	MAA_04103	MAC_03079
Signalosome subunit 5, regulation of sexual development	AN2129	NCU00467	OCS_01205	CCM_02582	BBA_04172	MAA_08893	MAC_05426
Transcription factor	AN6230	NCU07617	OCS_05103	CCM_01634	BBA_04959	MAA_05553	MAC_01182
Velvet activator VeA	AN1052	NCU01731	OCS_04577	CCM_04531	BBA_05501	MAA_01811	MAC_00039
Cytochrome FphA, required for sexual development	AN9008	NCU04834	OCS_06517	CCM_04461	BBA_02816	MAA_02298	MAC_04734

Table S6 Comparison of selected protein families between *O. sinensis* and other insect pathogens

Protein family	***O. sinensis***	***C. militaris***	***B. bassiana***	***M. roberstii***	***M. acridum***
Heterokaryon incompatibility proteins	5	15	21	34	22
Dehydrogenases	103	271	252	216	209
Cytochrome P450	57	57	83	123	100
Trypsins	2	12	23	32	17
Subtilisins	17	35	43	55	43
Aspartic proteases	12	21	21	33	25
Chitinases	9	20	20	21	14
Concanavalin A-like lectins	6	25	33	20	18
Glycoside hydrolases	66	135	145	159	130
Secondary metabolite biosynthesis backbone enzymes	28	30	42	61	40
Sugar/inositol transporters	10	3	6	5	5
Appressorium differentiation protein (MAS-like protein)	2	4	4	6	6
Adhesins	0	2	2	2	2
Chloroperoxidases	7	0	1	4	2

Table S7 Comparison of dehydrogenases encoded in the insect pathogens and plant obligate pathogen *B. graminis*

Enzyme family		EC No.	*O. sinensis*	*M. robertsii*	*M. acridum*	*B. bassiana*	*C. militaris*	*B. graminis*
Alcohol dehydrogenase	**Alcohol dehydrogenase**	**1.1.1.1**	**9**	**27**	**29**	**23**	**24**	**2**
	Alcohol dehydrogenase (NADP+)	**1.1.1.2**	**7**	**26**	**22**	**19**	**17**	**5**
D-xylulose reductase		1.1.1.9	2	1	1	1	0	0
Glycerate dehydrogenase		**1.1.1.29**	**0**	**4**	**4**	**1**	**1**	**0**
Glycerol-3-phosphate dehydrogenase(NAD+)		1.1.1.8	1	1	1	1	1	1
Glucose dehydrogenase	dTDP-4-Dehydrorhamnose reductase	1.1.1.133	1	1	1	1	1	0
	Glucose 1-dehydrogenase	**1.1.1.47**	**3**	**8**	**8**	**8**	**8**	**2**
	Glucose 6-phosphate sehydrogenase	1.1.1.49	1	1	1	1	1	1
	Glucose 1-dehydrogenase (NAD(+))	1.1.1.118	0	2	1	2	1	0
	Udp-glucose6-dehydrogenase	1.1.1.22	2	2	2	2	1	0
Homoserine dehydrogenase		1.1.1.3	1	1	1	1	1	1
Histidinol dehydrogenase		1.1.1.23	1	1	1	1	1	1
Sorbitol dehydrogenase (SDH)		1.1.1.14	5	7	7	4	8	1
Shikimate dehydrogenase		**1.1.1.25**	**2**	**4**	**5**	**5**	**5**	**1**
Lactate/Malate dehydrogenase	**3-isopropylmalate dehydrogenase**	**1.1.1.85**	**0**	**2**	**2**	**2**	**2**	**0**
	D-Lactate dehydrogenase	1.1.1.28	1	1	1	0	0	0
	L-Lactate dehydrogenase	1.1.1.27	0	1	2	1	1	0
	Malate dehydrogenase	1.1.1.37	2	2	2	2	2	2
	(R)-2-hydroxyacid dehydrogenase	1.1.1.272	0	1	0	1	1	0
3-hydroxybutyrate dehydrogenase		1.1.1.30	0	2	2	3	1	0
3-Hydroxyisobutyrate dehydrogenase		1.1.1.31	0	4	4	3	4	1
3-Hydroxyacyl-Coa dehydrogenase		1.1.1.35	5	8	7	6	6	4
Isocitrate dehydrogenase (NAD+)		1.1.1.41	1	2	2	2	2	2
Isocitrate dehydrogenase (NADP(+))		1.1.1.42	1	1	1	1	1	1
Phosphogluconate dehydrogenase		1.1.1.44	1	2	3	2	1	1
L-Gulonate 3-dehydrogenase		1.1.1.45	1	1	0	1	0	0
Hydroxysteroid Dehydrogenase (HSD)	3 Alpha-Hydroxysteroid dehydrogenase	1.1.1.50	0	2	2	1	0	1
	7-alpha-hydroxysteroid dehydrogenase	1.1.1.159	1	3	4	0	0	0
	11-beta-hydroxysteroid dehydrogenase	1.1.1.146	2	3	1	4	4	3
	estradiol 17-beta-dehydrogenase	1.1.1.62	**6**	**11**	**16**	**14**	**18**	**2**
	20-alpha-hydroxysteroid dehydrogenase	1.1.1.149	0	2	3	0	1	0
	3(or 17)-beta-hydroxysteroid dehydrogenase	1.1.1.51	0	2	0	0	0	0
	3-alpha-(or 20-beta)-hydroxysteroid dehydrogenase	1.1.1.53	3	6	5	3	3	0
Mannitol 2-dehydrogenase		1.1.1.67	0	0	0	0	0	1
Gluconate 5-dehydrogenase		**1.1.1.69**	**5**	**23**	**17**	**15**	**15**	**5**
Homoisocitrate dehydrogenase		1.1.1.87	0	1	1	1	1	1
Aryl-alcohol dehydrogenase		**1.1.1.90**	**0**	**3**	**2**	**1**	**2**	**0**
Phosphoglycerate dehydrogenase		**1.1.1.95**	**2**	**6**	**5**	**7**	**7**	**2**
L-Threonine-3-dehydrogenase		1.1.1.103	0	4	1	1	0	0
2-Deoxy-D-gluconate 3-dehydrogenase		1.1.1.125	0	0	1	0	0	0
GDP-Mannose 6-dehydrogenase		1.1.1.132	0	1	1	1	1	0

(*To be continued on the next page*)

(*Continued*)

Enzyme family		EC No.	*O. sinensis*	*M. robertsii*	*M. acridum*	*B. bassiana*	*C. militaris*	*B. graminis*
Mannitol dehydrogenase		1.1.1.138	0	2	3	4	3	1
15-Hydroxyprostaglandin dehydrogenase (NAD+)		**1.1.1.141**	**0**	**6**	**3**	**2**	**1**	**0**
IMP dehydrogenase		1.1.1.205	1	1	1	1	1	1
2-(R)-hydroxypropyl-CoM dehydrogenase		1.1.1.268	5	8	6	6	5	2
S-(hydroxymethyl)glutathione dehydrogenase		1.1.1.284	1	1	1	1	1	1
Formate Dehydrogenase		1.2.1.2	1	3	2	1	1	0
Aldehyde Dehydrogenase	**Aldehyde dehydrogenase (NAD+)**	**1.2.1.3**	**2**	**6**	**5**	**2**	**2**	**1**
	Aldehyde dehydrogenase (NAD(P)(+))	1.2.1.5	1	2	2	2	2	1
	Betaine-aldehyde dehydrogenase	1.2.1.8	1	2	4	1	1	1
	Aminobutyraldehyde dehydrogenase	1.2.1.19	0	1	1	0	0	0
	Glyceraldehyde-3-phosphate Dehydrogenase (NADP(+))	1.2.1.9	1	1	0	2	1	0
	Glyceraldehyde 3-phosphate dehydrogenase(phosphorylating)	1.2.1.12	1	1	1	1	1	1
	Lactaldehyde dehydrogenase	1.2.1.22	0	1	2	1	0	0
	Retinal dehydrogenase	**1.2.1.36**	**0**	**7**	**5**	**4**	**6**	**1**
Aspartate-semialdehyde dehydrogenase		1.2.1.11	0	1	1	1	1	1
Methylmalonate-semialdehyde dehydrogenase		1.2.1.27	1	1	1	1	1	1
N-acetyl-gamma-glutamyl-phosphate reductase		1.2.1.38	1	1	2	2	2	1
Glutamate 5-Semialdehyde		1.2.1.41	0	1	1	2	1	1
Formaldehyde dehydrogenase		**1.2.1.46**	**0**	**2**	**2**	**2**	**1**	**0**
5-Carboxymethyl-2-hydroxymuconate semialdehyde Dehydrogenase		1.2.1.60	0	1	1	1	1	1
2,3-Dihydro-2,3-dihydroxybenzoate dehydrogenase		1.3.1.28	1	1	2	1	2	0
15-Oxoprostaglandin 13-Oxidase		1.3.1.48	2	1	3	3	3	1
Cis-2,3-dihydrobiphenyl-2,3-diol dehydrogenase		1.3.1.56	0	2	1	0	0	0
Alanine dehydrogenase		1.4.1.1	1	1	1	1	1	1
Glutamic Dehydrogenase	Glutamate dehydrogenase (NAD(P+))	1.4.1.3	0	1	1	1	1	0
	Glutamate dehydrogenase (NADP(+))	1.4.1.4	0	1	1	1	1	0
Tetrahydrofolate dehydrogenase	Methylenetetrahydrofolate dehydrogenase (NADP(+))	1.5.1.5	1	2	2	1	0	1
	Formyltetrahydrofolate dehydrogenase	1.5.1.6	1	0	0	2	2	1
Saccharopine Dehydrogenase (NADP+, L-glutamate-forming)		1.5.1.10	1	1	1	1	1	1
1-Pyrroline-5-carboxylate dehydrogenase		1.5.1.12	0	2	1	2	1	1
Methylenetetrahydrofolate dehydrogenase (NAD(+))		1.5.1.15	1	1	1	1	1	0
Opine dehydrogenase		1.5.1.28	0	1	0	1	1	0
Proline dehydrogenase		1.5.99.8	0	1	1	2	1	0
NAD(P)H dehydrogenase (quinone)		1.6.5.2	0	1	1	1	1	0
NADPH dehydrogenase		**1.6.99.1**	**1**	**9**	**7**	**3**	**6**	**0**
NADH dehydrogenase (Quinone)		1.6.99.5	7	7	7	7	7	7
Dihydrolipoyl dehydrogenase		1.8.1.4	4	9	8	8	4	6
Xanthine dehydrogenase		1.17.1.4	1	1	1	1	1	1
Total			**103**	**271**	**252**	**216**	**209**	**76**

Protein families significantly under-represented (>2-fold) in *O. sinensis* are highlighted in bold.

Table S8 The core proteins involved in the biosynthesis of secondary metabolites in *O. sinensis* and other insect pathogens

Core enzyme[a]	*O. sinensis*	*C. militaris*	*B. bassiana*	*M. roberstii*	*M. acridum*
NRPS	6	7	14	11	10
NRPS-t1PKS	0	1	1	0	0
NRPS-Terpene	0	0	0	1	1
t1PKS	7	7	9	17	10
t1PKS-NRPS	1	5	5	8	5
t1PKS-Terpene	0	0	0	2	0
t1PKS-Terpene-NRPS	0	0	0	0	1
t3PKS	0	0	1	0	0
Terpene *synthase*	10	4	6	9	6
Terpene-NRPS	0	0	1	1	0
Other enzymes	6	6	5	12	7
Sum	**30**	**30**	**42**	**61**	**40**

[a]Abbreviations: NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; t1, type I; t3, type III. Other enzymes for either NPRS-like or PKS-like proteins.

Table S9 Comparison of proteins putatively involved in cold responses in *O. sinensis* and other insect pathogens

Function	Description	*O. sinensis*	*C. militaris*	*B. bassiana*	*M. roberstii*	*M. acridum*
Triacylglycerol accumulation	Aquaporin	1	1	1	1	1
	Glycerol kinase	2	1	1	2	2
	Phospholipid/glycerol acyltransferase	3	4	2	5	5
	Membrane bound O-acyl transferase	4	4	4	4	4
	Phosphatidic acid phosphatase type 2/haloperoxidase	4	6	6	6	6
	Diacylglycerol kinase	1	1	1	1	2
	Pyruvate kinase	4	7	5	7	6
	Perilipin-like protein	1	1	1	1	1
	Lipase	**22**	**38**	**43**	**47**	**41**
Fatty acid accumulation	**Fatty acid desaturase**	**7**	**3**	**3**	**4**	**2**
	Fatty acid synthase	1	1	1	2	2
	Fatty acid hydroxylase	**6**	**15**	**15**	**14**	**11**
Mannitol accumulation	Mannitol dehydrogenase	1	1	1	1	1
	Mannitol 1-phosphate 5-dehydrogenase, MPD	1	1	1	1	1
	Hexokinase, HXK	4	5	4	4	4
Trehalose accumulation	Trehalose-6-phosphate synthase/trehalose phosphatase (TPS/TPP)	2	3	5	2	2
	Trehalase	2	2	2	2	2
Heat response	Heat shock factor binding 1	2	2	3	2	1
	Heat shock protein HSP40	17	25	27	25	25
	Heat shock protein Hsp20	4	3	5	5	6
	Heat shock protein Hsp70	10	9	11	5	6
	Heat shock protein Hsp90	1	1	1	1	1
Cold response	**Insect-like antifreeze protein**	**2**	**0**	**0**	**0**	**0**
	Type II antifreeze protein	5	4	3	4	5
	RNA binding protein	1	1	1	2	2

Protein families significantly over- or under-represented in *O. sinensis* are highlighted in bold. Some bacteria and plants rely on antifreeze polyols such as trehalose and mannitol, but the *O. sinensis* genome encoded similar numbers of genes for these as other fungi.