Article

Not peer-reviewed version

Mutations and Differential Transcription of Mating-Type and Pheromone Receptor Genes in Hirsutella sinensis and Natural Cordyceps sinensis

Xiu-Zhang Li , Yu-Ling Li , Jia-Shi Zhu *

Posted Date: 29 March 2024

doi: 10.20944/preprints202403.1854.v1

Keywords: Differential transcription mating-type genes; mutant MAT1-1-1 protein; mutant α-pheromone receptor protein; genotypes of Ophiocordyceps sinensis; natural Cordyceps sinensis; transcriptional regulation; homothallism; pseudohomothallism; physiological heterothallism; hybridization

Article

Mutations and Differential Transcription of Mating-Type and Pheromone Receptor Genes in *Hirsutella sinensis* and Natural *Cordyceps sinensis*

Running Title: Differential Transcription of Mating-Type Genes in Hirsutella Sinensis

Xiu-Zhang Li, Yu-Ling Li and Jia-Shi Zhu *

State Key Laboratory of Plateau Ecology and Agriculture, Qinghai Academy of Animal and Veterinary Sciences, Qinghai University, Xining, Qinghai, China
* Correspondence: zhujosh@163.com

Abstract: The sexual reproduction of ascomycetes is controlled by the mating-type(*MAT*) locus. (Pseudo)homothallic reproduction has been hypothesized based on genetic data from *Hirsutella sinensis* (Genotype #1 of *Ophiocordyceps sinensis*). In this study, the differential occurrence of the mating-type genes in the *MAT1-1* and *MAT1-2* idiomorphs, which constitute the genetic control of *O. sinensis* reproduction, was analyzed in 237 strains and 5 genomes of *H. sinensis*. The mating-type genes were differentially transcribed in *H. sinensis* Strain L0106, and the MAT1-2-1 transcript in *H. sinensis* Strain 1229 was alternatively spliced, with an unspliced intron I containing stop codons; these findings constitute *O. sinensis* reproduction controls at the transcriptional and coupled transcriptional-translational levels. The occurrence of the mating-type genes and a-/α-pheromone receptor genes differed among the 5 *H. sinensis* genomes. The differential occurrence and transcription of these genes are inconsistent with self-fertilization under (pseudo)homothallism but suggest self-sterility of *H. sinensis*, which therefore needs a mating partner to achieve sexual outcrossing under heterothallism or hybridization. Consistent occurrence and transcription of the mating-type genes of both the *MAT1-1* and *MAT1-2* idiomorphs have been reported in natural and cultivated *Cordyceps sinensis*, but these findings are insufficient to determine the reproduction mode(s) because of the co-occurrence of >90 fungal species and multiple genomically independent genotypes of *O. sinensis*. The mutant MAT1-1-1 and α-pheromone receptor transcripts in natural *C. sinensis* result in the N-terminus or middle-truncated proteins, with changes in hydrophobicity and secondary structures of the proteins, suggesting heterogeneous fungal source(s) of the proteins and hybridization reproduction.

Keywords: Differential transcription mating-type genes; mutant MAT1-1-1 protein; mutant α-pheromone receptor protein; genotypes of *Ophiocordyceps sinensis*; natural *Cordyceps sinensis*; transcriptional regulation; homothallism; pseudohomothallism; physiological heterothallism; hybridization

Introduction

The sexual reproductive behavior of ascomycetes is controlled by transcription factors encoded at the mating-type (*MAT*) locus [Debuchy & Turgeon 2006; Jones & Bennett 2011; Zheng & Wang 2013; Wilson *et al*. 2015]. Hu *et al*. [2013] detected the MAT1-1-1 and MAT1-2-1 genes of the *MAT1-1* and *MAT1-2* idiomorphs in the genome assembly ANOV00000000 of *Hirsutella sinensis* Strain Co18 and hypothesized that *H. sinensis* (GC-biased Genotype #1 of *Ophiocordyceps sinensis*) experiences self-fertilization under homothallism. The MAT1-1-1 gene encodes a protein possessing a conserved α-box domain, and the MAT1-2-1 gene encodes a protein possessing a high-mobility-group (HMG) box domain [Zheng & Wang 2013; Wilson *et al*. 2015]. Bushley *et al*. [2013] described multicellular heterokaryotic structures of natural *Cordyceps sinensis* hyphae and ascospores containing mononucleated, binucleated, trinucleated, and tetranucleated cells; in addition, they detected the MAT1-1-1, MAT1-1-2 and MAT1-1-3 genes of the *MAT1-1* idiomorph and the MAT1-2-1 gene of the *MAT1-2* idiomorph, which are located >4 kb apart in the genome of *H. sinensis* Strain 1229. Based on these results, they hypothesized that *H. sinensis* underwent pseudohomothallism. However, Zhang *et al*. [2011] and Zhang and Zhang [2015] reported differential occurrence of the MAT1-1-1 and MAT1-2-1 genes in various *H. sinensis* strains or isolates and hypothesized that *O. sinensis* underwent facultative hybridization. These hypotheses regarding the sexual reproductive behavior of *O. sinensis* were proposed based on the genetic information of *H. sinensis* without considering the entire expression process of the *H. sinensis* mating-type genes, including genetic, epigenetic, transcriptional, posttranscriptional, translational and posttranslational modifications, as well as the activation and degradation of mating-type proteins.

Natural *C. sinensis* is one of most valued therapeutic agents in traditional Chinese medicine and has a rich history of clinical use in health maintenance, disease amelioration, postdisease and postsurgery recovery and antiaging therapy [Zhu *et al*. 1998a, 1998b, 2011]. The Chinese Pharmacopoeia defines natural *C. sinensis* as an insect-fungal complex containing the *O. sinensis* fruiting body and the remains of a Hepialidae moth larva (an intact, thick larval body wall with numerous bristles, intact larval intestine and head tissues, and fragments of other larval tissues) [Ren *et al*. 2013; Li *et al*. 2016b, 2022b; Zhu & Li 2017]. Studies of natural *C. sinensis* conducted using various histological, mycological and molecular approaches have demonstrated its multicellular heterokaryotic structure and genetic heterogeneity, as revealed by the detection of at least 17 genomically independent genotypes of *O. sinensis* [Kinjo & Zang 2001; Stensrud *et al*. 2007; Xiao *et al*. 2009; Zhang *et al*. 2009; Zhu *et al*. 2010, 2012; Chen *et al*. 2011; Li *et al*. 2013, 2016b, 2022b, 2023a, 2023b; Mao *et al*. 2013] and >90 fungal species spanning at least 37 genera and larval genes [Qiu *et al*. 1987; Li *et al*. 1988, 2016b, 2019, 2022b; Dai *et al*. 1989; Engh 1999; Jiang & Yao 2003; Chen *et al*. 2004; Stensrud *et al*. 2005; Leung *et al*. 2006; Zhu *et al*. 2007, 2023b; Yang *et al*. 2008; Zhang *et al*. 2010; Barseghyan *et al*. 2011; Xia *et al*. 2015; Wang *et al*. 2020]. The expression of mating-type genes during the sexual life of natural *C. sinensis* is much more complex than that in pure fungal cultures and is mutually and/or antagonistically disrupted by the expression of the metagenomes of multiple cocolonized fungi and by larval host innate immunity and acquired immunological responses during *O. sinensis* fungal infection and proliferation [Meng *et al*. 2015; Lu *et al*. 2016; Wang *et al*. 2016, 2018; Li *et al*. 2019, 2022b]. It has been suggested that larval tissues are not just culture medium that passively provide nutrients for fungal growth, similar to the medium used in *in vitro* culture or fermentation [Lu *et al*. 2016; Wang *et al*. 2018; Li *et al*. 2022b]. Notably, the Latin name *Cordyceps sinensis* has been indiscriminately used since the 1840s for both the teleomorph/holomorph of *C. sinensis* fungus and the wild insect-fungal complex, and the fungus was renamed *O. sinensis* in 2007 [Sung *et al*. 2007; Ren *et al*. 2013; Zhang *et al*. 2012; Zhu & Wu 2015; Li *et al*. 2016b, 2022b; Zhu & Li 2017]. In this paper, we refer to the fungus/fungi as *O. sinensis* and continue the customary use of the name *C. sinensis* to refer to the wild or cultivated insect-fungi complex, although this practice will likely be replaced by the discriminative use of exclusive Latin names in the future.

Many studies on natural *C. sinensis* and *O. sinensis* fungi have focused primarily on *H. sinensis*, which has been postulated to be the sole anamorph of *O. sinensis* [Wei *et al*. 2006]. Following this sole anamorph hypothesis, *H. sinensis* Strain EFCC7287, Genotype #1 of *O. sinensis*, was selected for use as a reference species for the renaming of *Cordyceps sinensis* to *Ophiocordyceps sinensis* [Sung *et al*. 2007]. However, this renaming did not involve the natural insect-fungal complex and did not cover the remaining 16 genotypes of *O. sinensis* fungi, for which pure cultures have not been available for taxonomy and nomenclature multigene studies or for genomics, transcriptomics, proteomics, natural chemistry and pharmacology studies [Engh 1999; Kinjo & Zang 2001; Stensrud *et al*. 2005, 2007; Xiao *et al*. 2009; Zhang *et al*. 2009, 2012; Zhu *et al*. 2010, 2012, 2018; Chen *et al*. 2011; Gao *et al*. 2011, 2012; Yao *et al*. 2011; Li *et al*. 2013, 2016b, 2016c, 2020, 2022b, 2023a, 2023b]. Numerous studies have reported that the sequences of *O. sinensis* Genotypes #2–17, regardless of whether they are GC or AT biased,

doi:10.20944/preprints202403.1854.v1

are not found in the genome assemblies ANOV00000000, JAAVMX000000000, LKHE00000000, LWBQ00000000, and NGJJ00000000 of the GC-biased Genotype #1 *H. sinensis* Strains Co18, IOZ07, 1229, ZJB12195, and CC1406-203, respectively [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020], providing evidence of the genomic independence of the 17 *O. sinensis* genotypes, which are interindividual *O. sinensis* fungi [Stensrud *et al*. 2007; Xiao *et al*. 2009; Zhang *et al*. 2009; Zhu *et al*. 2010, 2012; Chen *et al*. 2011; Li *et al*. 2016b, 2016c, 2020, 2021, 2022a, 2022b, 2023a, 2023b]. In this paper, we continue using the anamorphic name *H. sinensis* for Genotype #1 of *O. sinensis*, although a group of mycologists [Zhang *et al*. 2013] have improperly implemented the “One Fungus=One Name” nomenclature rule of the International Mycological Association while disregarding the presence of multiple genomically independent genotypes of *O. sinensis* fungi and inappropriately replacing the anamorphic name *H. sinensis* with the teleomorphic name *O. sinensis* [Ren *et al*. 2013; Zhu & Wu 2015; Zhu & Li 2017; Zhu *et al*. 2018; Li *et al*. 2022b].

The lifecycle of *C. sinensis* includes asexual and sexual growth stages during its development and maturation. *O. sinensis* infection of the larvae of the Hepialidae family initiates the lifecycle of natural *C. sinensis*, which includes several developmental phases: (1) formation of a stromal primordium; (2) immature *C. sinensis* with a short stroma (1–2 cm) (*O. sinensis* grows asexually in the first two phases); (3) maturing *C. sinensis* in the transition from asexual to sexual growth with an approximately 3–5 cm stroma without the formation of an expanded fertile portion close to the tip of the stroma; and (4) mature *C. sinensis* that grows sexually with a long stroma (usually >5 cm) with an expanded fertile portion close to the stromal tip, densely covering numerous ascocarps before and after ascospore ejection (Figure 1).

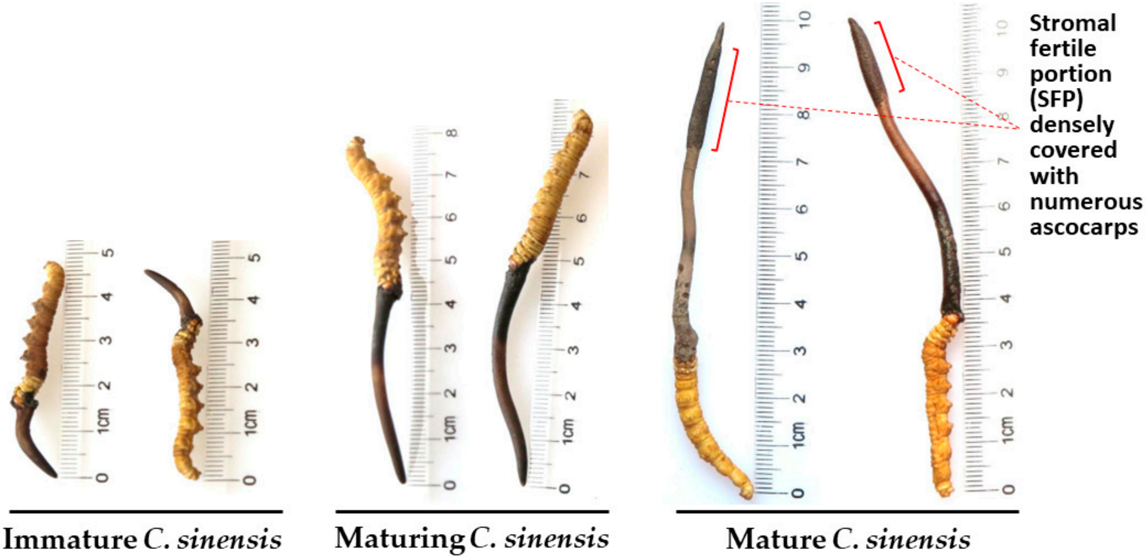


Figure 1. Images of immature, maturing, and mature natural *C. sinensis* (modified from [Zhu *et al*. 2010; Yao *et al*. 2011, 2014]). The right two panels show mature *C. sinensis* specimens with the expanded stromal fertile portion (SFP) close to the tip of the *C. sinensis* stroma, which is densely covered with numerous ascocarps.

C. sinensis maturation is associated with dramatic changes in the mycobiota, metagenome, metatranscriptomic and proteomic expression, chemical constituent fingerprint and pharmacological functions [Zhu *et al*. 2007, 2010, 2012; Yang *et al*. 2008; Liang *et al*. 2010; Gao *et al*. 2011, 2012; Dong *et al*. 2014; Yao *et al*. 2014; Zhu & Wu 2015; Li *et al*. 2016b, 2016c, 2020, 2022b, 2023b; Zhu & LI 2017]. Li *et al*. [2023a, 2023b] reported naturally and semiejected *C. sinensis* ascospores associated with differential co-occurrence of several GC- and AT-biased genotypes of *O. sinensis*. The biomasses/abundances of GC- and AT-biased genotypes of *O. sinensis* change dynamically in an asynchronous, disproportional manner in the caterpillar body, stroma, stromal fertile portion (with ascocarps) and ascospores of natural *C. sinensis* during maturation [Xiao *et al*. 2009; Zhu *et al*. 2010, 2012; Gao *et al*. 2011, 2012; Yao *et al*. 2011; Zhu & Wu 2015; Li *et al* 2016b, 2023a, 2023b; Zhu & Li 2017].

Thus, it is essential to identify the maturation stages of natural *C. sinensis* specimens during the examination of mating-type gene expression.

Multiple genetic, genomic and transcriptomic sequences of numerous *H. sinensis* strains and natural *C. sinensis* specimens are available in the GenBank database, which enables further examination of the hypotheses regarding the sexual reproduction strategy of *O. sinensis* at the genome and transcriptome levels. This paper discusses the genetic occurrence of mating-type genes in 237 *H. sinensis* strains, the transcription of mating-type genes and pheromone receptor genes involved in *O. sinensis* sexual reproduction in *H. sinensis* and natural *C. sinensis*, and other features related to the potential mating process of *O. sinensis*.

Materials and Methods

Gene and Genome Sequences of H. sinensis Strains and Natural C. sinensis

Table 1 lists 237 *H. sinensis* strains that were used to obtain mating-type gene sequences, which are available in GenBank and the original papers [Zhang *et al*. 2009, 2011; Bushley *et al*. 2013; Hu *et al*. 2013; Liu *et al*. 2015, 2020; Zhang & Zhang 2015; Li *et al*. 2016a, 2021, 2022a; Jin *et al*. 2020; Shu *et al*. 2020]. Five sets of genome assemblies, ANOV00000000, JAAVMX000000000, LKHE00000000, LWBQ00000000, and NGJJ00000000, of *H. sinensis* Strains Co18, IOZ07, 1229, ZJB12195 and CC1406-20395, respectively, are available in GenBank [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020].

Table 1. Differential occurrence of the MAT1-1-1 and MAT1-2-1 genes in 237 *H. sinensis* strains.

22 strains containing only the MAT1-1-1 gene	**66 strains containing only the MAT1-2-1 gene**	**149 strains containing both the MAT1-1-1 and MAT1-2-1 genes**		
CS09-143	CC1406-203XZ-CD-41	1229NP10_1	SC09_47	XZ09_32
CS09-229	CS2XZ-CD-59	Co18NP10_2	SC09_57	XZ09_46
GS03	GS09_337XZ-CD-64	CS6-251QH01	SC09_65	XZ09_48
IOZ07	QH02XZ-ML-191	CS18-266QH03	SC09_77	XZ09_59
QH07_188	QH05XZ-LZ05-6	CS25-273QH04	SC09_87	XZ09_71
QH07_197	QH09_11XZ-LZ06_1	CS26-277QH06	SC09_97	XZ09_80
XZ05_3	QH09_187XZ-LZ06_7	CS34-291QH07	SC09_107	XZ09_95
XZ05_6	QH-LJ-214XZ-LZ06_21	CS36-1294QH08	SC09_117	XZ09_100
XZ07_H2	QH-LJ-236XZ-LZ06_61	CS37-295QH09_20L	SC09_128	XZ09_106
XZ07_108	QH-QL-206XZ-LZ06_108	CS68-2-1228QH09_33L	SC09_147	XZ09_113
XZ07_154	QH-QL-207XZ-LZ07_30	CS68-2-1229QH09_37	SC09_157	XZ09_118
XZ07_166	QH-YS-188XZ-LZ07_64	CS68-5-1216QH09_46	SC09_167	XZ10_7
XZ07_176	QH-YS-189XZ-LZ07_H1	CS70-1208QH09_56	SC09_180	XZ10_15
XZ07_180	QH-YS-196XZ-LZ07_H2	CS70-1211QH09_66	SC09_190	XZ10_17
XZ08_4	QH-YS-197AXZ-NQ_74	CS70-1212QH09_78	SC09_200	XZ10_23
XZ08_10	QH-YS-197BXZ-NQ_80	CS71-1218QH09_93	SC10_4	XZ12_1
XZ08_26	QH-YS-199XZ-NQ_84	CS71-1219QH09_122	SC10_18	XZ12_16
XZ08_59	SC-2XZ-NQ_86	CS71-1220QH09_131	SC10_21	XZ12_33
XZ08_A1	SC-3XZ-NQ_92	CS76-1284QH09_151	TB01	XZ12_43
XZ08_B1	SC-4XZ-NQ_139	CS91-1291QH09_164	TB02	YN01
YN07_6	SC-5XZ-NQ_154	CS560-961QH09_173	TB03	YN02
YN07_8	SC-7XZ-NQ_155	CS561-964QH09_201	TB04	YN03
	SCK05-4-3XZ-NQ_156	GS01QH09_210	TB05	YN09_3
	XZ07_11XZ-NQ_166	GS02QH10_1	TB06	YN09_6
	XZ07_46XZ-NQ_176	GS04QH10_4	TB07	YN09_22
	XZ08_33XZ-NQ_180	GS05QH10_7	TB08	YN09_51
	XZ08_38XZ-SN_44	GS09_111SC01	XZ05_2	YN09_61

XZ-CD-A1YN-1	GS09_121SC02	XZ05_7	YN09_64
XZ-CD-B1YN-4	GS09_131SC03	XZ05_8	YN09_72
XZ-CD-4YN-5	GS09_143SC04	XZ05_12	YN09_81
XZ-CD-10YN-6	GS09_201SC05	XZ06_124	YN09_85
XZ-CD-26YN-8	GS09_225SC06	XZ06_152	YN09_89
XZ-CD-30ZJB12195	GS09_229SC07	XZ06_260	YN09_96
	GS09_281SC08	XZ07_133	YN09_101
	GS09_311SC09_1	XZ08_24	YN09_140
	GS10_1SC09_21	XZ08_56	
	GS10_10SC09_36	XZ09_4	
	ID10_1SC09_37	XZ09_15	

Genomic DNA from *H. sinensis* Strain Co18 was sequenced with the Roche 454 GS FLX system (Illumina HiSeq: 454), and the shotgun sequences were assembled using SOAPdenovo v.1.05 and Newbler v.2.3 [Hu *et al*. 2013]. Genomic DNA from Strain 1229 was sequenced with Illumina HiSeq sequencing technology, and the shotgun sequences were assembled using ABySS v.1.2.3 [Li *et al*. 2016a]. Genomic DNA from Strain ZJB12195 was sequenced with Illumina sequencing technology (HiSeq 2000 Sequencing System), and the shotgun sequences were assembled using SOAPdenovo v.2.0 [Jin *et al*. 2020]. Genomic DNA from Strain CC1406-203 was sequenced with continuous long reads (CLRs) following the Hierarchical Genome Assembly Process (HGAP) workflow (PacBioDevNet; Pacific Biosciences) and assembled with CA software (v.7.0) and the PacBio Rs_PreAssembler.1 module [Liu *et al*. 2020]. Genomic DNA from Strain IOZ07 was sequenced with PacBio Sequel sequencing technology, and the shotgun sequences were assembled using Canu v.1.7 [Shu *et al*. 2020].

Transcriptome and Metatranscriptome Assemblies and Transcripts of Mating-Type Genes of H. sinensis Strains and Natural C. sinensis

The transcriptome assembly GCQL00000000 is accessible in GenBank and was derived from fermented mycelia of the *H. sinensis* Strain L0106 [Liu *et al*. 2015]. The mycelia were collected for total RNA extraction from cultures grown for 3, 6 and 9 days. Total RNA (20 mg per sample) was subjected to mRNA purification, and total mRNA was used to construct a cDNA library and sequenced using Illumina HiSeq sequencing technology. The shotgun nucleotide sequences were assembled using SOAPdenovo v.2.0.

Natural *C. sinensis* specimens (unknown maturational status) were collected from Kangding County, Sichuan Province, China [Xiang *et al*. 2015]. Total RNA from this specimen was sequenced using 454 technology. Sequences longer than 50 bp from the 454 reads were assembled using GS De Novo Assembler software v 2.6 or Newbler 2.6 (454 Life Sciences Corporation, USA). The shotgun sequences were assembled as a metatranscriptome assembly under GenBank accession number GAGW00000000 using Newbler v.2.3 and 2.6.

Another metatranscriptome assembly was derived from specimens of mature natural *C. sinensis* that were collected from Deqin, Yunnan Province, China (*cf*. the Appendix of [Xia *et al*. 2017]). Total RNA was extracted from the fruiting bodies to construct a cDNA library, which was sequenced using the Illumina HiSeq2000 platform. The raw reads were assembled using Trinity (version r20140717) [Grabherr *et al*. 2011]. The sequencing reads were deposited in GenBank under BioProject #PRJNA382001. According to the Appendix of [Xia *et al*. 2017], the metatranscriptome assembly sequences were uploaded to the depository database, www.plantkingdomgdb.com/Ophiocordyceps_sinensis/data/cds/Ophiocordyceps_sinensis_CDS.fas, which is currently inaccessible, but a previously downloaded cDNA file was used for the metatranscriptomic analysis.

Sequence Alignment Analysis

All gene, genome, transcript, transcriptome and metatranscriptome sequences and other PCR-amplified DNA sequences of the mating-type and pheromone receptor genes of *H. sinensis* and natural *C. sinensis* were reanalyzed using the MegaBlast, discontinuous MegaBlast, Blastp, Tblastn, or blastn suite-SRA programs provided by GenBank (https://blast.ncbi.nlm.nih.gov/) to compare and align nucleotide and amino acid sequences.

Amino Acid Property and Scale Analysis

The hydrophobicity/hydrophilicity of the mating-type and pheromone receptor proteins of *H. sinensis* and natural *C. sinensis* were characterized based on the chemical-physical properties of the component amino acids using Peptide 2.0 (https://www.peptide2.com/N_peptide_hydrophobicity_hydrophilicity.php). The component amino acids were scaled (Table 2) and plotted at a window size of 21 amino acid residues for the α-Helix, β-Sheet, β-Turn and Coil of the proteins using the linear weight variation model of the ExPASy PlotScale algorithm (https://web.expasy.org/protscale/) provided by the SIB Swiss Institute of Bioinformatics [Chou & Fasman 1978; Deleage & Roux 1987; Gasteiger *et al*. 2005]. The amino acid hydrophobicity properties and the waveforms of PlotScale plots for the proteins were compared to explore potential alterations in the folding and secondary structures of the mating-type and pheromone receptor proteins.

Table 2. Amino acid properties and scale: Conformational parameters for α-Helix, β-Sheet, and Coil [Deleage & Roux 1987] and β-Sheet [Chou & Fasman 1978].

			Chemical-physical property	**α-Helix**	**β-Sheet**	**β-Turn**	**Coil**
Aspartic acid	Asp	D	Acidic	0.924	0.540	1.197	1.197
Glutamic acid	Glu	E	Acidic	1.504	0.370	1.149	0.761
Alanine	Ala	A	Aliphatic	1.489	0.830	0.788	0.824
Isoleucine	Ile	I	Aliphatic	1.003	1.600	0.240	0.886
Leucine	Leu	L	Aliphatic	1.236	1.300	0.670	0.810
Valine	Val	V	Aliphatic	0.990	1.700	0.387	0.772
Phenylalanine	Phe	F	Aromatic	1.195	1.380	0.624	0.797
Tryptophan	Trp	W	Aromatic	1.090	1.370	0.546	0.941
Tyrosine	Tyr	Y	Aromatic	0.787	1.470	0.795	1.109
Arginine	Arg	R	Basic	1.224	0.930	0.912	0.893
Histidine	His	H	Basic	1.003	0.870	0.970	1.068
Lysine	Lys	K	Basic	1.172	0.740	1.302	0.897
Asparagine	Asn	N	with polar neutral side chains	0.772	0.890	1.572	1.167
Cysteine	Cys	C	with polar neutral side chains	0.966	1.190	0.965	0.953
Glutamine	Gln	Q	with polar neutral side chains	1.164	1.100	0.997	0.947
Methionine	Met	M	with polar neutral side chains	1.363	1.050	0.436	0.810
Serine	Ser	S	with polar neutral side chains	0.739	0.750	1.316	1.130
Threonine	Thr	T	with polar neutral side chains	0.785	1.190	0.739	1.148
Glycine	Gly	G	Unique amino acids	0.510	0.750	1.860	1.251
Proline	Pro	P	Unique amino acids	0.492	0.550	1.415	1.540

Results

Differential Occurrence of Mating-Type Genes in H. sinensis Strains

A search of the GenBank database and the literature revealed intraspecific variations in 237 *H. sinensis* strains that reportedly contain the MAT1-1-1 and MAT1-2-1 genes (Table 1): 22 strains contain only the MAT1-1-1 gene, 66 strains harbor only the MAT1-2-1 gene, and 149 strains harbor both genes [Zhang *et al*. 2011, 2013; Bushley *et al*. 2013; Hu *et al*. 2013; Zhang & Zhang 2015; Li *et al*. 2016a, 2021,

doi:10.20944/preprints202403.1854.v1

2022a; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020]. These findings provide evidence of intraspecific variations in *H. sinensis* strains [Li *et al*. 2022a].

Bushley *et al*. [2013] identified 3 mating-type genes of the *MAT1-1* idiomorph in the genome sequence KC437356 of *H. sinensis* Strain 1229: the MAT1-1-1 gene (6,530→7,748), MAT1-1-2 gene (4,683→6,183) and MAT1-1-3 gene (3,730→4,432). The sequences of these mating-type genes from KC437356 show 99.9–100% homology with the segment sequences of the genome assemblies LKHE01001116 (3,691←4,909; 5,374←6,874; 7,125←7,827) of Strain 1229; ANOV01017390 (302←1,519) and ANOV01017391 (276←1,776; 2,027←2,729) of Strain Co18; and JAAVMX010000001 (6,698,911→6,700,129; 6,696,939→6,698,439; 6,695,986→6,696,688) of Strain IOZ07 (Table 3) [Bushley *et al*. 2013; Hu *et al*. 2013; Li *et al*. 2016a; Shu *et al*. 2020]. Note that the arrows "→" and "←" indicate sequences in the sense and antisense chains of the genomes, respectively. However, the mating-type genes of the *MAT1-1* idiomorph are absent from the genome assemblies NGJJ00000000 and LWBQ00000000 of *H. sinensis* Strains CC1406-203 and ZJB12195, respectively [Jin *et al*. 2020; Liu *et al*. 2020]. A BLAST search of the GenBank database identified the MAT1-1-1 gene in 127 *H. sinensis* strains with 98.0–100% homology among the sequences but did not find any sequences of the MAT1-1-2 or MAT1-1-3 genes in these strains [Zhang *et al*. 2009, 2013; Bushley *et al*. 2013; Hu *et al*. 2013; Li *et al*. 2021, 2022b].

The MAT1-2-1 sequence JQ325153 of *H. sinensis* Strain GS09_121 [Zhang *et al*. 2014] exhibits 99.7–99.9% homology with the segment sequences of the genome assemblies LKHE01001605 (13,851←14,723), ANOV01000063 (9,319→10,191), NGJJ01000619 (23,021←23,893) and LWBQ01000021 (238,864←239,736) of *H. sinensis* Strains 1229, Co18, CC1406-203 and ZJB12195, respectively (Table 3) [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Liu *et al*. 2020]. However, the MAT1-2-1 gene is absent in the genome assembly JAAVMX000000000 of Strain IOZ07 [Shu *et al*. 2020]. Further searches of the GenBank database identified the MAT1-2-1 gene sequence in 183 *H. sinensis* strains with 97.3–100% homology [Zhang *et al*. 2009, 2011, 2013, 2014; Bushley *et al*. 2013; Hu *et al*. 2013; Zhang & Zhang 2015; Li *et al*. 2021].

Table 3. Percent similarity between the mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs and the genome and transcriptome sequences of *H. sinensis* and natural *C. sinensis*.

	MAT1-1-1	**MAT1-1-2**	**MAT1-1-3**	**MAT1-2-1**
	(*vs*. KC437356 of Strain CS68-2-1229)			(*vs*. JQ325153 of Strain GS09_121)
	% similarity of the genome sequences			
H. sinensis Strain 1229	100%	100%	100%	99.9%
H. sinensis Strain Co18	99.9%	100%	100%	99.7%
H. sinensis Strain IOZ07	100%	100%	100%	—
H. sinensis Strain CC1406-203	—	—	—	99.9%
H. sinensis Strain ZJB12195	—	—	—	99.9%
	% similarity of the coding sequences of the genes (excluding introns) and the transcriptome sequences			
H. sinensis Strain L0106	—	—	—	99.0%
H. sinensis Strain 1229	100%	100%	100%	100%
Natural *C. sinensis* (Deqin, Yunnan)	100%	—	—	100%
Natural *C. sinensis* (Kangding, Sichuan)	100%	—	—	—

Bushley *et al*. [2013] found the presence of monokaryotic, bikaryotic, trikaryotic and tetrakaryotic structures in multicellular hyphae and ascospores and showed that the *MAT1-1* and *MAT1-2* idiomorphs are not closely linked because the MAT1-1-1 and MAT1-2-1 genes are located more than 4 kb apart in the *H. sinensis* genome. They then proposed that *H. sinensis* undergoes pseudohomothallism. The MAT1-1-1 and MAT1-2-1 genes are located on contig_17390 (302←1,519) of ANOV01017390 and contig_63 (9,319→10,191) of ANOV01000063 in the genome of Strain Co18, and scaffold_1116 (3,691←4,909) of LKHE01001116 and scaffold_1605 (13,851←14,723) of LKHE01001605 in the genome of Strain 1229 [Hu *et al*. 2013; Li *et al*. 2016a]. The MAT1-1-1 gene is present in the genome assembly JAAVMX010000001 of Strain IOZ07, but the MAT1-2-1 gene is absent [Shu *et al*. 2020]. The MAT1-2-1 gene is present in the genome assemblies NGJJ01000619 and LWBQ01000021 of Strains CC1406-203 and ZJB12195, respectively, but the MAT1-1-1 gene is absent [Jin *et al*. 2020; Liu *et al*. 2020].

Zhang *et al*. [2011] cloned the MAT1-2-1 gene but not the mating-type genes within the *MAT1-1* idiomorph in *H. sinensis* Strains CS2 and SCK05-4-3. The differential occurrence of mating-type genes was confirmed by genomic evidence in the *H. sinensis* strains CC1406-203, IOZ07 and ZJB12195 (Table 3) [Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020]. It is believed that the homothallic reproduction of ascomycetes requires the combined expression of the mating-type genes of both *MAT1-1* and *MAT1-2* idiomorphs within a single fungal cell and that the expression of only one of the MAT1-1-1 and MAT1-2-1 genes might reciprocally require the expression of complementary mating-type genes from another fungus of the same species for heterothallic outcrossing or from a different species for hybridization [Bushley *et al*. 2013; Zheng & Wang 2013; Wilson *et al*. 2015; Zhang *et al*. 2018]. Although the homothallism and pseudohomothallism hypotheses were proposed for *H. sinensis* based on the occurrence of the mating-type genes of both *MAT1-1* and *MAT1-2* idiomorphs in *H. sinensis* Strains Co18 and 1229 [Bushley *et al*. 2013; Hu *et al*. 2013], Zhang and Zhang [2015] reported the differential occurrence of mating-type genes of the *MAT1-1* and/or *MAT1-2* idiomorphs in 125 wild-type isolates of natural *C. sinensis* collected from different production areas in geographically remote locations and hypothesized that *H. sinensis* undergoes facultative hybridization.

Notably, Zhang and Zhang [2015] reported 4.9% and 6.1% allelic variation (93.9% and 95.1% similarity) in the full-length sequences of the MAT1-1-1 and MAT1-2-1 genes, respectively, and 4.7% and 5.7% variation in the exon sequences of numerous *C. sinensis* isolates. These variations might disrupt the translation of the coding sequences of the genes. However, a BLAST search of the GenBank database revealed much less variation (97.3–100% similarity) among 127 sequences of the MAT1-1-1 gene and 183 sequences of the MAT1-2-1 gene of *H. sinensis* strains and *C. sinensis* isolates, suggesting that Zhang and Zhang [2015] may not have uploaded all analyzed mating-type gene sequences to GenBank.

Differential Transcription of Mating-Type Genes in the Transcriptome of H. sinensis Strains

In addition to the mating-type genes in *H. sinensis*, transcription of the mating-type genes of both the *MAT1-1* and *MAT1-2* idiomorphs and the production and full activation of the encoded proteins within a single cell are needed to accomplish the homothallic mating process and to control the development and maturation of fruiting bodies, ascocarps and ascospores [Turgeon & Yoder 2000; Debuchy & Turgeo 2006; Jones & Bennett 2011; Zheng & Wang 2013; Wilson *et al*. 2015]. The sequences of the MAT1-1-1, MAT1-1-2 and MAT1-1-3 genes of *H. sinensis* are absent in the transcriptome assembly GCQL00000000 of *H. sinensis* Strain L0106 (Table 3) [Liu *et al*. 2015]. Zhang *et al*. [2011] found no transcripts of the mating-type genes within the *MAT1-1* idiomorph in *H. sinensis* strains and wild-type *C. sinensis* isolates. However, Bushley *et al*. [2013] obtained the transcript cDNA of the MAT1-1-1 gene with the splicing of two introns in the mycelia of Strain 1229. Therefore, it is apparent that the genes of the *MAT1-1* idiomorph are differentially transcribed in *H. sinensis* Strains 1229, CS2, L0106, and SCK05-4-3 and many other strains in which the genes of the *MAT1-1* idiomorph are absent in their genomes (*cf*. Table 1).

The MAT1-2-1 gene JQ325153 of *H. sinensis* GS09_121 contains two introns: “Intron Phase-Two” for intron I (52 nucleotides [nt]; 187→238 of JQ325153) and “Intron Phase-One” for intron II (55 nt; 535→589 of JQ325153). Both intron phases disrupt the open reading frame and codons (Figure 2) [Zhang *et al*. 2011, 2014; Bushley *et al*. 2013; Zhang & Zhang 2015; Li *et al*. 2021]. Figure 2 also shows 3 other MAT1-2-1 genes (FJ654187, JQ325237, and KM197536) from *H. sinensis* Strains CS2, XZ12_16, and XZ12_16, respectively, which were 97.7–99.5% similar to the sequence of JQ325153. A single T-to-C transition downstream of the third stop codon was present in intron I in the sequences of JQ325237 and KM197536 and apparently had no impact on the three stop codons in intron I.

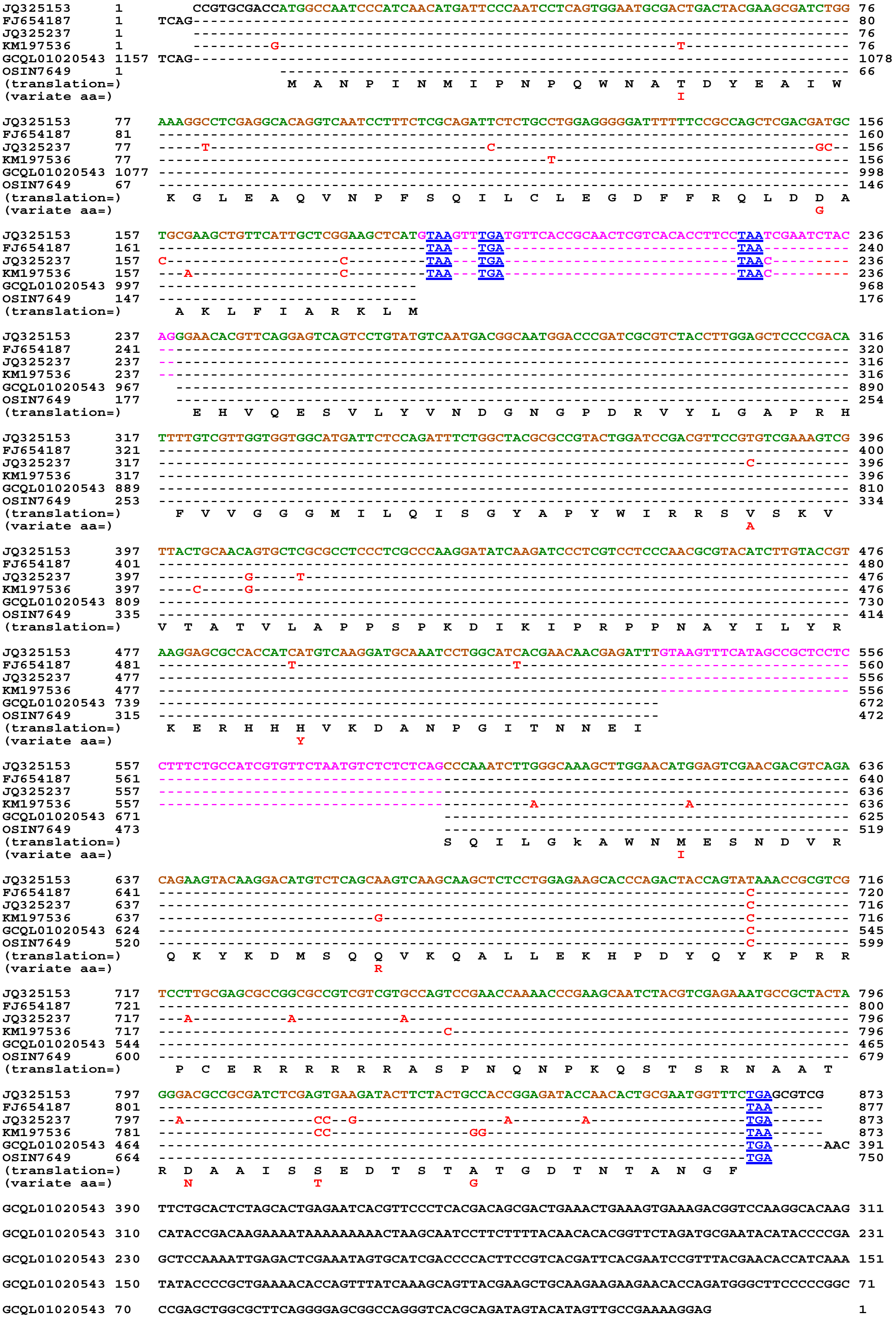


Figure 2. Alignment of the MAT1-2-1 gene sequences with the transcriptome sequences obtained from *H. sinensis* Strain L0106 and natural *C. sinensis*. JQ325153, FJ654187, JQ325237 and KM197536 are the MAT1-2-1 gene sequences were found in *H. sinensis* Strains GS09_121, CS2, XZ12_16, and XZ12_16, respectively [Zhang *et al*. 2009, 2014]. GCQL01020543 is the transcriptome sequence obtained from *H. sinensis* Strain L0106 [Liu *et al*. 2015]. OSIN7649 is the metatranscriptome sequence obtained from

natural *C. sinensis* collected from Deqin County, Yunnan Province, China [Xia *et al*. 2017]. The triplets shown in alternating brown and green indicate the open reading frame. The sequences in pink represent introns, and the underlined triplets "TAA" and "TGA" shown in blue are stop codons. The protein sequence was translated from the transcript sequences. The nucleotide and amino acid residues in red represent point mutations and variants, respectively. Hyphens indicate identical bases, and spaces denote unmatched sequence gaps.

The coding sequences of 3 exons (388←671; 672←967; 968←1,153) of the MAT1-2-1 gene JQ325153 exhibit 99.7–100% homology with the transcriptome assembly GCQL01020543 of *H. sinensis* Strain L0106 [Zhang *et al*. 2014; Liu *et al*. 2015]. Xia *et al*. [2017] confirmed that the MAT1-2-1 transcript sequence (OSIN7649) with the splicing of both introns I and II in the metatranscriptome assembly of natural *C. sinensis* (further discussion below). Bushley *et al*. [2013] reported that "comparison of DNA and cDNA sequences of MAT1-2-1 revealed only one spliced intron of 55 bp" in *H. sinensis* Strain 1229, which "did not result from contamination of genomic DNA" and illustrated the alternatively spliced MAT1-2-1 cDNA with unspliced intron I and spliced intron II. Figure 2 herein shows the MAT1-2-1 sequences, which possess 3 stop codons, "TAA" or "TGA", within intron I (the first, third and 13th triplets), leading to translational arrest of the MAT1-2-1 transcript with unspliced intron I.

To rule out the possibility that gene mutations within MAT1-2-1 intron I occur to avoid translational interruption due to the presence of stop codons within intron I, a BLAST search of the GenBank database was performed. This search revealed that the intron I sequence of the MAT1-2-1 gene is 100% identical to the MAT1-2-1 gene segments of the genome assemblies LKHE01001605, NGJJ01000619, ANOV01000063 and LWBQ01000021 of Strains 1229, CC1406-203, Co18 and ZJB12195, respectively, and to the DNA sequences of more than 100 *H. sinensis* strains. A five-base insertion occurs in intron I of the MAT1-2-1 genes FJ654204 and FJ654205 in two *H. sinensis* isolates, XZ-LZ07-H1 and XZ-LZ07-H2, obtained from the Nyingchi District in Tibet, which was postulated to be the center of origin of the species *H. sinensis* [Zhang *et al*. 2009]. The insertion is located downstream of the first stop codon "TAA". Thus, the alternatively spliced MAT1-2-1 gene transcript with unspliced intron I might represent an additional mechanism of the coupled transcriptional-translational regulation of the mating process of *H. sinensis* Strain 1229 [Bushley *et al*. 2013], distinct from the gene silencing used for transcriptional control (Table 3), although the full expression process (genetic, epigenetic, transcriptional, posttranscriptional, translational, and posttranslational modifications; protein activation-degradation; etc.) of the *H. sinensis* MAT1-2-1 gene needs to be thoroughly examined in future reproductive physiology studies.

The above transcriptome analysis of *H. sinensis* Strains 1229 and L0106 [Bushley *et al*. 2013; Liu *et al*. 2015] revealed the transcriptional and coupled transcriptional-translational controls of mating-type gene expression in *H. sinensis*. These results are inconsistent with the hypotheses of self-fertilization under homothallism and pseudohomothallism for *H. sinensis* proposed by Bushley *et al*. [2013] and Hu *et al*. [2013]. Instead, these findings suggest that *O. sinensis* requires sexual partners from monoecious or dioecious *H. sinensis* for physiological heterothallic reproduction or heterospecific fungal species to ensure hybridization if the species are able to overcome interspecific reproductive isolation.

The available evidence regarding mating-type gene transcription might explain why previous efforts to cultivate pure *H. sinensis* in research-oriented academic settings to induce the production of fruiting bodies and ascospores have consistently failed [Holliday & Cleaver 2008; Stone 2010; Hu *et al*. 2013; Zhang *et al*. 2013] and why a successful inoculation-cultivation project in a product-oriented industrial setting presented a species contradiction between the GC-biased Genotype #1 *H. sinensis* strains used as anamorphic inoculants and the teleomorphic AT-biased Genotype #4 of *O. sinensis* in the fruiting body of cultivated *C. sinensis,* as reported by Wei *et al*. [2016].

Differential Occurrence of Mating-Type Genes in Natural and Cultivated C. sinensis Insect-Fungi Complexes

The above sections discuss the differential occurrence and transcription of mating-type genes in pure *H. sinensis* strains (GC-biased Genotype #1 of *O. sinensis*) and wild-type *C. sinensis* isolates. Mycobiota studies have demonstrated the differential coexistence of >90 cocolonized fungi belonging to at least 37 fungal genera in the stroma and caterpillar body of natural *C. sinensis* [Zhang *et al*. 2010; Xia *et al*. 2015], which has been confirmed by metagenomic and metatranscriptomic studies and by the identification of >100 mitogenomic transcript repeats in natural and cultivated *C. sinensis* [Xiang *et al*. 2014; Meng *et al*. 2015; Xia *et al*. 2017; Wang *et al*. 2018; Zhang *et al*. 2018; Zhong *et al*. 2018; Li *et al*. 2019; Zhao *et al*. 2020; Yang *et al*. 2021]. Many studies have reported differential co-occurrence of several GC- and AT-biased genotypes of *O. sinensis* in different compartments (the stroma, caterpillar body, stromal fertile portion densely covered with numerous ascocarps, and ascospores) of natural *C. sinensis*, the sequences of which reside in the independent fungal genomes of different *O. sinensis* fungi [Engh 1999; Kinjo & Zang 2001; Chen *et al*. 2011; Stensrud *et al*. 2005, 2007; Xiao *et al*. 2009; Zhang *et al*. 2009; Zhu *et al*. 2010, 2012, 2018, 2019; Li *et al*. 2013; 2016b, 2016c, 2020, 2022b, 2023a, 2023b; Mao *et al*. 2013; Wei *et al*. 2016; Zhu & Li 2017]. Notably, the indiscriminate use of the Latin name *Ophiocordyceps sinensis* for multiple genotypes of *O. sinensis* fungi and natural and cultivated insect-fungal complexes has prompted scientists to carefully distinguish the study materials when composing and reading research papers/reports to ensure correct wording and understanding [Li *et al*. 2016b, 2020, 2022a, 2022b, 2023a, 2023b; Zhu & Li 2017].

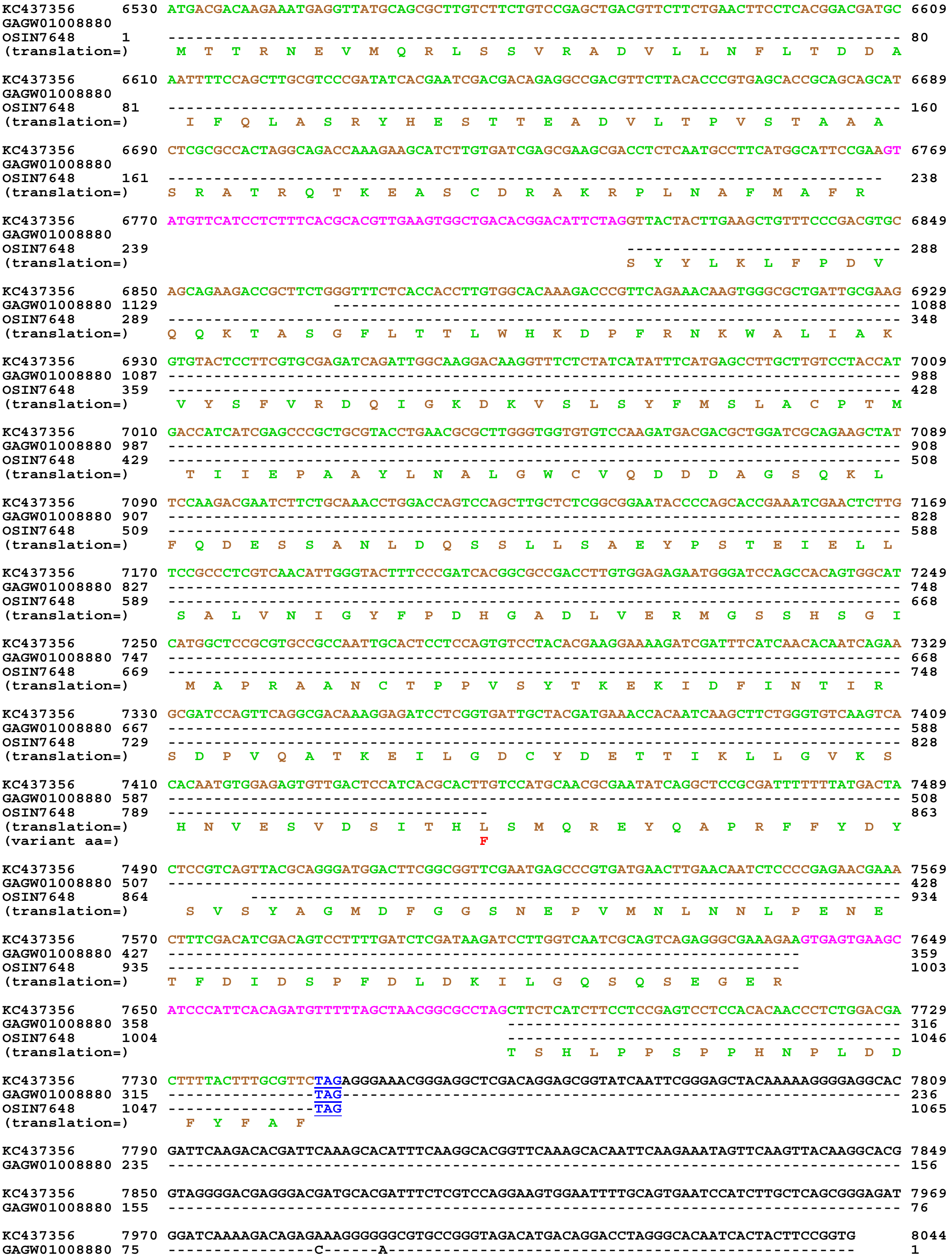
The mating-type genes of both the *MAT1-1* and *MAT1-2* idiomorphs reportedly exist in natural and cultivated *C. sinensis*, for instance, (1) in the metagenome of mature natural *C. sinensis* [Xia *et al*. 2017]; (2) in the compartments of natural *C. sinensis* at different developmental stages [Zhong *et al*. 2018]; (3) in cultivated *C. sinensis* at different developmental phases [Li *et al*. 2019]; and (4) in the early-developed stroma and caterpillar body of natural *C. sinensis* with very low read count values and in 31 other *C. sinensis* specimens [Zhao *et al*. 2020]. However, the detection of the mating-type genes of both the *MAT1-1* and *MAT1-2* idiomorphs reported in these studies might not correctly represent the genetic-based capability for self-fertilization of a single *H. sinensis* cell due to the differential coexistence of multiple cocolonized fungi and multiple genotypes of *O. sinensis* in the compartments of natural and cultivated *C. sinensis* [Engh 1999; Kinjo & Zang 2001; Xiao *et al*. 2009; Zhang *et al*. 2009; Zhu *et al*. 2010, 2012; Gao *et al*. 2011, 2012; Chen *et al*. 2011; Yao *et al*. 2011; Li *et al*. 2013, 2016b, 2020, 2022, 2023a, 2023b; Mao *et al*. 2013; Zhang & Zhang 2015; Wei *et al*. 2016; Zhu & Li 2017; Zhong *et al*. 2018; Zhao *et al*. 2020; Yang *et al*. 2021]. Hence, the possibility of physiological heterothallic outcrossing or hybridization should not be neglected, particularly because of the species contradiction between the anamorphic inoculants of GC-biased *H. sinensis* strains used in industrial cultivation projects and the sole teleomorph of AT-biased Genotype #4 in cultivated *C. sinensis* (Wei *et al*. 2016).

Differential Transcription of Mating-Type Genes in Natural and Cultivated C. sinensis Insect-Fungi Complexes

Transcription of Mating-Type Genes in the Metatranscriptome Assemblies of Natural *C. sinensis*

Xiang *et al*. [2014] reported a metatranscriptome project with natural *C. sinensis* specimens (unknown maturation stage) collected from Kangding County of Sichuan Province, China. The assembled metatranscriptome segment 297←1,129 of GAGW01008880 aligns to the MAT1-1-1 gene sequence KC437356 but is shorter due to the absence of exon I, intron I and 48 bp of exon II of the MAT1-1-1 gene (Figure 3). Its intron II (48 bp) is spliced between transcript nucleotides 358 and 359. Excluding the missing segments, GAGW01008880 is 100% identical to the remaining coding sequences (771 bp of exon II and the full length of exon III) of the MAT1-1-1 gene.

Xia *et al*. [2017] reported a metagenome-metatranscriptome project with fully mature *C. sinensis* specimens collected from Deqin County of Yunnan Province, China. The metatranscriptome assembly OSIN7648 in the Appendix [Xia *et al*. 2017] exhibits 87.3% similarity to the MAT1-1-1 gene sequence KC437356 and possesses 3 spliced/deleted segments, introns I and II (52 bp and 48 bp in pink in Figure 3, respectively), and an additional deleted segment of 54 bp (in blue in Figure 3) within exon II. Excluding the 3 spliced/deleted segments, OSIN7648 shares 100% identity with the coding sequences of exons I and III and the remaining exon II.



```
KC437356     6530 ATGACGACAAGAAATGAGGTTATGCAGCGCTTGTCTTCTGTCCGAGCTGACGTTCTTCTGAACTTCCTCACGGACGATGC 6609
GAGW01008880
OSIN7648     1    -------------------------------------------------------------------------------- 80
(translation=)     M  T  T  R  N  E  V  M  Q  R  L  S  S  V  R  A  D  V  L  L  N  F  L  T  D  D  A

KC437356     6610 AATTTTCCAGCTTGCGTCCCGATATCACGAATCGACGACAGAGGCCGACGTTCTTACACCCGTGAGCACCGCAGCAGCAT 6689
GAGW01008880
OSIN7648     81   -------------------------------------------------------------------------------- 160
(translation=)      I  F  Q  L  A  S  R  Y  H  E  S  T  T  E  A  D  V  L  T  P  V  S  T  A  A  A

KC437356     6690 CTCGCGCCACTAGGCAGACCAAAGAAGCATCTTGTGATCGAGCGAAGCGACCTCTCAATGCCTTCATGGCATTCCGAAGT 6769
GAGW01008880
OSIN7648     161  ------------------------------------------------------------------------------   238
(translation=)    S  R  A  T  R  Q  T  K  E  A  S  C  D  R  A  K  R  P  L  N  A  F  M  A  F  R

KC437356     6770 ATGTTCATCCTCTTTCACGCACGTTGAAGTGGCTGACACGGACATTCTAGGTTACTACTTGAAGCTGTTTCCCGACGTGC 6849
GAGW01008880
OSIN7648     239                                                    ------------------------------ 288
(translation=)                                                      S  Y  Y  L  K  L  F  P  D  V

KC437356     6850 AGCAGAAGACCGCTTCTGGGTTTCTCACCACCTTGTGGCACAAAGACCCGTTCAGAAACAAGTGGGCGCTGATTGCGAAG 6929
GAGW01008880 1129                   -------------------------------------------------------------- 1088
OSIN7648     289  -------------------------------------------------------------------------------- 348
(translation=)    Q  Q  K  T  A  S  G  F  L  T  T  L  W  H  K  D  P  F  R  N  K  W  A  L  I  A  K

KC437356     6930 GTGTACTCCTTCGTGCGAGATCAGATTGGCAAGGACAAGGTTTCTCTATCATATTTCATGAGCCTTGCTTGTCCTACCAT 7009
GAGW01008880 1087 -------------------------------------------------------------------------------- 988
OSIN7648     359  -------------------------------------------------------------------------------- 428
(translation=)     V  Y  S  F  V  R  D  Q  I  G  K  D  K  V  S  L  S  Y  F  M  S  L  A  C  P  T  M

KC437356     7010 GACCATCATCGAGCCCGCTGCGTACCTGAACGCGCTTGGGTGGTGTGTCCAAGATGACGACGCTGGATCGCAGAAGCTAT 7089
GAGW01008880 987  -------------------------------------------------------------------------------- 908
OSIN7648     429  -------------------------------------------------------------------------------- 508
(translation=)      T  I  I  E  P  A  A  Y  L  N  A  L  G  W  C  V  Q  D  D  D  A  G  S  Q  K  L

KC437356     7090 TCCAAGACGAATCTTCTGCAAACCTGGACCAGTCCAGCTTGCTCTCGGCGGAATACCCCAGCACCGAAATCGAACTCTTG 7169
GAGW01008880 907  -------------------------------------------------------------------------------- 828
OSIN7648     509  -------------------------------------------------------------------------------- 588
(translation=)    F  Q  D  E  S  S  A  N  L  D  Q  S  S  L  L  S  A  E  Y  P  S  T  E  I  E  L  L

KC437356     7170 TCCGCCCTCGTCAACATTGGGTACTTTCCCGATCACGGCGCCGACCTTGTGGAGAGAATGGGATCCAGCCACAGTGGCAT 7249
GAGW01008880 827  -------------------------------------------------------------------------------- 748
OSIN7648     589  -------------------------------------------------------------------------------- 668
(translation=)     S  A  L  V  N  I  G  Y  F  P  D  H  G  A  D  L  V  E  R  M  G  S  S  H  S  G  I

KC437356     7250 CATGGCTCCGCGTGCCGCCAATTGCACTCCTCCAGTGTCCTACACGAAGGAAAAGATCGATTTCATCAACACAATCAGAA 7329
GAGW01008880 747  -------------------------------------------------------------------------------- 668
OSIN7648     669  -------------------------------------------------------------------------------- 748
(translation=)      M  A  P  R  A  A  N  C  T  P  P  V  S  Y  T  K  E  K  I  D  F  I  N  T  I  R

KC437356     7330 GCGATCCAGTTCAGGCGACAAAGGAGATCCTCGGTGATTGCTACGATGAAACCACAATCAAGCTTCTGGGTGTCAAGTCA 7409
GAGW01008880 667  -------------------------------------------------------------------------------- 588
OSIN7648     729  -------------------------------------------------------------------------------- 828
(translation=)    S  D  P  V  Q  A  T  K  E  I  L  G  D  C  Y  D  E  T  T  I  K  L  L  G  V  K  S

KC437356     7410 CACAATGTGGAGAGTGTTGACTCCATCACGCACTTGTCCATGCAACGCGAATATCAGGCTCCGCGATTTTTTTATGACTA 7489
GAGW01008880 587  -------------------------------------------------------------------------------- 508
OSIN7648     789  -----------------------------------                                              863
(translation=)     H  N  V  E  S  V  D  S  I  T  H  L  S  M  Q  R  E  Y  Q  A  P  R  F  F  Y  D  Y
(variant aa=)                                       F

KC437356     7490 CTCCGTCAGTTACGCAGGGATGGACTTCGGCGGTTCGAATGAGCCCGTGATGAACTTGAACAATCTCCCCGAGAACGAAA 7569
GAGW01008880 507  -------------------------------------------------------------------------------- 428
OSIN7648     864           ----------------------------------------------------------------------- 934
(translation=)      S  V  S  Y  A  G  M  D  F  G  G  S  N  E  P  V  M  N  L  N  N  L  P  E  N  E

KC437356     7570 CTTTCGACATCGACAGTCCTTTTGATCTCGATAAGATCCTTGGTCAATCGCAGTCAGAGGGCGAAAGAAGTGAGTGAAGC 7649
GAGW01008880 427  ---------------------------------------------------------------------            359
OSIN7648     935  ---------------------------------------------------------------------            1003
(translation=)    T  F  D  I  D  S  P  F  D  L  D  K  I  L  G  Q  S  Q  S  E  G  E  R

KC437356     7650 ATCCCATTCACAGATGTTTTTAGCTAACGGCGCCTAGCTTCTCATCTTCCTCCGAGTCCTCCACACAACCCTCTGGACGA 7729
GAGW01008880 358                                       ------------------------------------------- 316
OSIN7648     1004                                      ------------------------------------------- 1046
(translation=)                                         T  S  H  L  P  P  S  P  P  H  N  P  L  D  D

KC437356     7730 CTTTTACTTTGCGTTCTAGAGGGAAACGGAGGCTCGACAGGAGCGGTATCAATTCGGGAGCTACAAAAAGGGGAGGCAC 7809
GAGW01008880 315  ----------------TAG------------------------------------------------------------- 236
OSIN7648     1047 ----------------TAG                                                              1065
(translation=)      F  Y  F  A  F

KC437356     7790 GATTCAAGACACGATTCAAAGCACATTTCAAGGCACGGTTCAAAGCACAATTCAAGAAATAGTTCAAGTTACAAGGCACG 7849
GAGW01008880 235  -------------------------------------------------------------------------------- 156

KC437356     7850 GTAGGGGACGAGGGACGATGCACGATTTCTCGTCCAGGAAGTGGAATTTTGCAGTGAATCCATCTTGCTCAGCGGGAGAT 7969
GAGW01008880 155  -------------------------------------------------------------------------------- 76

KC437356     7970 GGATCAAAAGACAGAGAAAGGGGGGCGTGCCGGGTAGACATGACAGGACCTAGGGCACAATCACTACTTCCGGTG      8044
GAGW01008880 75   ----------------C------A---------------------------------------------------      1
```

Figure 3. Alignment of the MAT1-1-1 genome sequences FC437356 and KR004003 of *H. sinensis* Strains CS68-2-1229 and GS09_225 and the transcriptome sequences of natural *C. sinensis*. FC437356 is the MAT1-1-1 gene of *H. sinensis* Strain CS68-2-1229 [2013]. The sequence segments in pink represent introns I and II. The metatranscriptome assemblies OSIN7648 and GAGW01008880 are the MAT1-1-

1 transcripts of natural *C. sinensis* specimens collected from Deqin County of Yunnan Province and Kangding County of Sichuan Province, China, respectively [Xiang *et al*. 2014; Xia *et al*. 2017]. The triplets shown in alternating brown and green indicate the open reading frame. The underlined "TAG" triplets in blue represent stop codons. The protein sequence was translated from the transcript sequences. Amino acid residues in red represent variants. Hyphens indicate identical bases, and spaces denote unmatched sequence gaps.

MAT1-1-1 transcripts in both metatranscriptome assemblies of natural *C. sinensis* are short compared with the coding sequence (1,119 bp) of the MAT1-1-1 gene of KC437356 [Bushley *et al*. 2013; Xiang *et al*. 2014; Xia *et al*. 2017]. As described above, OSIN7648 (1,065 bp) contains an additional deleted segment (54 bp in blue in Figure 3) within exon II of the MAT1-1-1 gene [Xia *et al*. 2017], which may produce a middle-truncated MAT1-1-1 protein that is missing 18 aa residues. Compared with the MAT1-1-1 gene, GAGW01008880 is 286 bp shorter due to the absence of a 238-bp segment of exon I and a 48-bp segment of exon II [Xiang *et al*. 2014] and may produce an N-terminal-truncated MAT1-1-1 protein. Figure 3 also shows a 296-bp noncoding segment of GAGW01008880 flanking downstream of the stop codon "TAG".

The transcript sequences of the MAT1-1-2 and MAT1-1-3 genes were not identified in the metatranscriptome assemblies of the natural *C. sinensis* specimens [Xiang *et al*. 2014; Xia *et al*. 2017].

The MAT1-2-1 gene sequence was not detected in the metatranscriptome assembly GAGW00000000 of natural *C. sinensis* collected from Kangding County of Sichuan Province (*cf*. Figure 2) [Xiang *et al*. 2014]. However, the MAT1-2-1 gene sequence was found in the metatranscriptome assembly OSIN7649 of natural *C. sinensis* collected from Deqin County of Yunnan Province, which exhibits 100% homology with the transcriptome assembly GCQL01020543 of *H. sinensis* Strain L0106 [Liu *et al*. 2015; Xia *et al*. 2017].

Transcription of Mating-Type Genes in Unassembled Metatranscriptome Sequencing Read Archives of Natural and Cultivated *C. sinensis*

Zhang *et al*. [2018] reported differential transcription of the mating-type genes of natural *C. sinensis* at the early maturing stage. The MAT1-1-1 gene (OCS_06642) was reportedly transcribed at a similar level in all five specimens collected from the counties of Qumalai (specimen #1), Gonghe (specimen #2) and Zaduo (specimen #3) of Qinghai Province and Naqu (specimen #4) and Nyingchi (specimen #5) in Tibet. However, the heatmap shown in Figure 3B of [Zhang *et al*. 2018] demonstrated very different transcription in all five specimens. The MAT1-1-3 gene (OCS_06644) was transcribed at a low level only in specimen #4 but was silent in the other specimens, whereas the MAT1-2-1 gene (OCS_00196) was transcribed at low levels in specimens #4 and #5 but was silent in specimens #1–3.

Zhong *et al*. [2018] observed differential transcription of mating-type genes in the early-developed stroma and caterpillar body of natural *C. sinensis* collected from Yushu, Qinghai Province, China. They reported that "the MAT1-1-1 gene could be detected in all three groups of samples" but did not report whether the transcript or cDNA of the MAT1-1-1 gene was detected in parallel. The authors also reported that "MAT1-1-2 was not detected in all samples". They identified the transcripts of the MAT1-1-3 and MAT1-2-1 genes at "a very low read count value" but were unsure whether they "might participate in fruiting body development" in natural *C. sinensis*. Thus, they encouraged further studies aiming to verify the transcription of mating-type genes and the functions of mating proteins in natural *C. sinensis* at different developmental stages.

Li *et al*. [2019] observed the differential transcription of the mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs at very low levels (with read counts of 0–40 and fragments per kilobase of exon model per million mapped fragments [FPKM] values of 0–8.4) in cultivated *C. sinensis* at all developmental phases (an FPKM value less than 10 indicates very low levels of transcription). The mating-type genes of the *MAT1-1* idiomorph were not transcribed in the hyphae (read count=0) but were differentially transcribed at very low levels in cultivated *C. sinensis* at other developmental stages. The MAT1-2-1 gene was transcribed at very low levels (FPKM 0.16–0.93) in cultivated *C. sinensis* at all developmental stages.

Zhao *et al*. [2020] reported nearly no transcription of the mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs (transcripts per million reads [TPM] of 0–2.27). The authors concluded that these genes might not play any role in the initiation of fruiting body development in cultivated *C. sinensis*.

We conducted a BLAST search of the GenBank database against the unassembled transcript reads as part of (1) the metatranscriptome BioProject PRJNA325365 and the Sequence Read Archive (SRA) SRP076425 of natural *C. sinensis* [Zhang *et al*. 2018]; (2) the SRA SRP103894 of natural *C. sinensis* at an early developmental stage [Zhong *et al*. 2018]; (3) the BioProject GSE123085 of cultivated *C. sinensis* at all developmental phases [Li *et al*. 2019]; and (4) the BioProject PRJNA600609 of cultivated *C. sinensis* at the initiation stage [Zhao *et al*. 2020]. We identified hundreds to thousands of unassembled transcript reads of MAT1-1-1, MAT1-1-2 and MAT1-2-1 cDNA. These unassembled transcript reads share 73.8–100% similarity to the coding sequences of the *H. sinensis* MAT1-1-1, MAT1-1-2 and MAT1-2-1 genes. No match was found between sequences of the *H. sinensis* MAT1-1-3 gene and the unassembled transcript reads of natural and cultivated *C. sinensis*. This range of variation between the sequences of the *H. sinensis* mating-type genes and the unassembled transcript reads indicates multiple fungal sources of the transcripts in natural and cultivated *C. sinensis*.

This section discusses the coexistence of multiple mating-type genes at the metatranscriptome level in natural and cultivated *C. sinensis* [Xiang *et al*. 2014; Xia *et al*. 2017; Zhang *et al*. 2018; Zhong *et al*. 2018; Li *et al*. 2019; Zhao *et al*. 2020]. The differential transcription of the mating-type genes of both the *MAT1-1* and *MAT1-2* idiomorphs in natural and cultivated *C. sinensis*, which contain >90 fungal species spanning at least 37 genera and multiple GC- and AT-biased genotypes of *O. sinensis*, provides insufficient evidence to prove that *H. sinensis* is the source of the transcripts and therefore does not sufficiently support the hypothesis that *H. sinensis* (Genotype #1 of *O. sinensis*) uses homothallic/pseudohomothallic mating strategies [Zhang *et al*. 2011, 2013, 2014, 2018; Bushley *et al*. 2013; Hu *et al*. 2013; Xiang *et al*. 2014; Zhang & Zhang 2015; Xia *et al*. 2017; Zhong *et al*. 2018; Li *et al*. 2019, 2022b, 2023c, 2024; Zhao *et al*. 2020].

Variations in Mating Proteins

The above sections explored the differential occurrence and transcription of the mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs at the genome/metagenome and transcriptome/metatranscriptome levels and discussed the genetic, transcriptional, and coupled transcriptional-translational controls of the mating processes of *H. sinensis* in natural and cultivated *C. sinensis*. Allelic mutations in mating-type genes and their transcripts might influence protein translation and posttranslational processes.

The MAT1-1-1 protein sequences derived from the gene sequences of 127 *H. sinensis* strains show 98.1–100% nucleotide homology [Zhang *et al*. 2009, 2011, 2013, 2014; Bushley *et al*. 2013; Hu *et al*. 2013; Zhang & Zhang 2015]. Among these, 29 sequences contained one to seven conserved and/or nonconserved mutated residues of a total of 372 aa residues at single-variant sites (Figure 4), sharing 98.1–99.7% similarity with the MAT1-1-1 protein sequence AGW27560 of Strain CS68-2-1229. The MAT1-1-1 protein sequence (354 aa) translated from the metatranscriptome assembly OSIN7648 obtained from natural *C. sinensis* specimens collected from Deqin County, Yunnan Province, China, shows 92.5–94.9% similarity to the MAT1-1-1 protein sequence (372 aa) of *H. sinensis*, with a substitution of leucine (L)→phenylalanine (F) and a deletion of 18 aa (SMQREYQAPREFYDYSVS) (Figures 3 and 4) [Zhang *et al*. 2011, 2014; Bushley *et al*. 2013; Hu *et al*. 2013; Zhang & Zhang 2015; Xia *et al*. 2017]. This mutation likely produces a middle-truncated MAT1-1-1 protein. The MAT1-1-1 protein (276 aa) translated from the metatranscriptome assembly GAGW01008880 obtained from natural *C. sinensis* specimens collected from Kangding County of Sichuan Province, China, is 96 aa shorter (MTTRNEVMQRLSSVRADVLLNFLTDDAIFQLASRY-HESTTEADVLTPVSTAAASRATRQTKEASCDRAKRPLNAFMAFRSYYLKLFPDVQQQKTASG) at its N-terminus (Figures 3 and 4) [Zhang *et al*. 2011, 2014; Bushley *et al*. 2013; Hu *et al*. 2013; Xiang *et al*. 2014; Zhang & Zhang 2015]. This mutation likely produces an N-terminal-truncated MAT1-1-1 protein.

```
AGW27560  1   MTTRNEVMQRLSSVRADVLLNFLTDDAIFQLASRYHESTTEADVLTPVSTAAASRATRQTKEASCDRAKRPLNAFMAFRSYYLKLFPDVQ 90
OSIN7648      ------------------------------------------------------------------------------------------
GAGW01008880
ALH24995  1   ------------------------------------------------------------------------------------------ 90
ALH24951  1   ------------------------------------------------------------------------------------------ 90
ALH24992  1   --------------------------------------------------T--------------------------------------- 90
ALH25047  1   ------------------------------------------------------------------------------------------ 90
ALH24948  1   ------------------------------------------------------------------------------------------ 90
ALH25003  1   ------------------------------------------------------------------------------------------ 90
ALH25005  1   ------------------------------------------------------------------------------------------ 90
ALH25057  1   ------------T----I--------------------------------V----T---------------------------------- 90
ALH25001  1   ------------T----I--------------------------------V----T---------------------------------- 90
ALH25048  1   ------------------------------------------------------------------------------------------ 90
ALH24999  1   ------------T----I--------------------------------V----T---------------------------------- 90

AGW27560  91  QKTASGFLTTLWHKDPFRNKWALIAKVYSFVRDQIGKDKVSLSYFMSLACPTMTIIEPAAYLNALGWCVQDDDAGSQKLFQDESSANLDQ 180
OSIN7648      ------------------------------------------------------------------------------------------
GAGW01008880        ------------------------------------------------------------------------------------
ALH24995  91  ------------------------------------------------------------------------------------------ 180
ALH24951  91  ------------------------------------------------------------------------------------------ 180
ALH24992  91  ------------------------------------------------------------------------------------------ 180
ALH25047  91  ------------------------------------------------------------------------------------------ 180
ALH24948  91  ---------------------------------SF------------------------------------------------------- 180
ALH25003  91  ------------------------------------------------------------------------------------------ 180
ALH25005  91  ------------------------------------------------------------------------------------------ 180
ALH25057  91  ------------------------------------------------------------------------------------------ 180
ALH25001  91  ------------------------------------------------------------------------------------------ 180
ALH25048  91  ------------------------------------------------------------------------------------------ 180
ALH24999  91  ------------------------------------------------------------------------------------------ 180

AGW27560  181 SSLLSAEYPSTEIELLSALVNIGYFPDHGADLVERMGSSHSGIMAPRAANCTPPVSYTKEKIDFINTIRSDPVQATKEILGDCYDETTIK 270
OSIN7648      ------------------------------------------------------------------------------------------
GAGW01008880  ------------------------------------------------------------------------------------------
ALH24995  181 ------------------------------------------------------------------------------------------ 270
ALH24951  181 ------------------------------------------------------------------------------------------ 270
ALH24992  181 ------------------------------------------------------------------------------------------ 270
ALH25047  181 ------------------------------------------------------------------------------------------ 270
ALH24948  181 ------------------------------------------------------------------------------------------ 270
ALH25003  181 --------------------------------------G--------------------------------------------------- 270
ALH25005  181 -------------------------------------------------------------V---------------------------- 270
ALH25057  181 -------------------------------------------------------------V---------------------------- 270
ALH25001  181 -------------------------------------------------------------------------------------K---- 270
ALH25048  181 ----------------------------------I----------T--------------------I----------------------- 270
ALH24999  181 -------------------------------------------------------------V---------------------------- 270

AGW27560  241 LLGVKSHNVESVDSITHLSMQREYQAPRFFYDYSVSYAGMDFGGSNEPVMNLNNLPENETFDIDSPFDLDKILGQSQSEGERTSHLPPSP 360
OSIN7648      -----------------F                  ------------------------------------------------------
GAGW01008880  ------------------------------------------------------------------------------------------
ALH24995  241 ------------------------------------------------------------------------------------------ 360
ALH24951  241 -------------------------------------V---------------------------------------------------- 360
ALH24992  241 ------------------------------------------------------------------------------------------ 360
ALH25047  241 -----------------------------------------------------------------------------L------------ 360
ALH24948  241 ------------------------------------------------------------------------------------------ 360
ALH25003  241 ------------------------------------------------------------------------------------------ 360
ALH25005  241 ------------------------------------------------------------------------------------------ 360
ALH25057  241 -------------------------------------------------------H---------------------------------- 360
ALH25001  241 -----------------------H------------------------------------------------------------------ 360
ALH25048  241 ------------------------------------------A----------------------------------------------- 360
ALH24999  241 -------------------------------------------------------H---------------------------------- 360

AGW27560  361 PHNPLDDFYFAF                                                                               372
OSIN7648      ------------
GAGW01008880  ------------
ALH24995  361 ------------                                                                               372
ALH24951  361 ------------                                                                               372
ALH24992  361 ------------                                                                               372
ALH25047  361 ------------                                                                               372
ALH24948  361 ------------                                                                               372
ALH25003  361 ------------                                                                               372
ALH25005  361 ------------                                                                               372
ALH25057  361 -Y----------                                                                               372
ALH25001  361 ------------                                                                               372
ALH25048  361 ------------                                                                               372
ALH24999  361 -Y----------                                                                               372
```

Figure 4. Alignment of the MAT1-1-1 protein sequences translated from the transcripts of *H. sinensis* strains and metatranscriptomes of natural *C. sinensis*. OSIN7648 and GAGW01008880 are MAT1-1-1 protein sequences derived from metatranscriptome sequences of natural *C. sinensis* collected from Deqin County of Yunnan Province and Kangding County of Sichuan Province in China, respectively [Xiang *et al*. 2014; Xia *et al*. 2017]. AGW27560, ALH24995, ALH24951, ALH24992, ALH25047, ALH24948, ALH25003, ALH25005, ALH25057, ALH25001, ALH25048 and ALH24999 are the MAT1-1-1 protein sequences of Strains 1229, SC09_97, GS09_229, SC09_65, YN09_61, GS09_143, XZ05_6, XZ05_8, XZ12_16, XZ05_2, YN09_64 and XZ07_H2, respectively [Bushley *et al*. 2013; Zhang & Zhang 2015]. The residues in red indicate the variant amino acids. Hyphens indicate identical amino acid residues, and spaces denote unmatched sequence gaps.

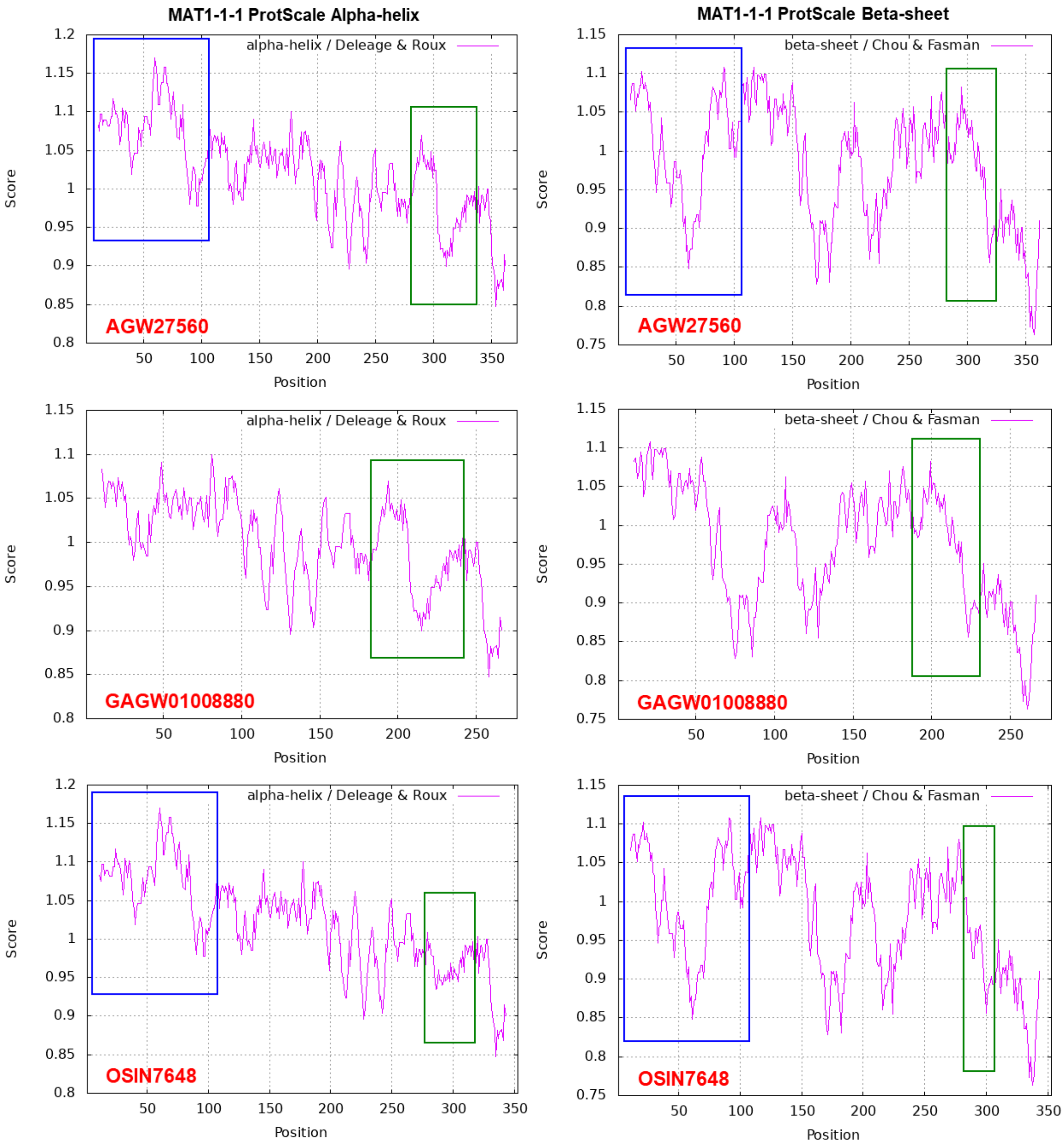
The absence of the MAT1-1-1 transcript in the transcriptome of *H. sinensis* Strain L0106 and the remarkable variations in the MAT1-1-1 protein conformation in natural *C. sinensis* as translated from the MAT1-1-1 transcripts in the metatranscriptome assemblies may suggest heterogeneous fungal sources of the protein in natural insect-fungal complexes (Table 3) [Xiang *et al*. 2014; Liu *et al*. 2015; Xia *et al*. 2017].

Table 4 lists the 4 types of amino acids in MAT1-1-1 proteins based on their chemical-physical and hydrophobic properties determined using Peptide 2.0 [Peters & Elofsson 2014; Simm *et al*. 2016]. Compared with the percentage of amino acids of the MAT1-1-1 protein AGW27560 of *H. sinensis* Strain CS68-2-1229, changes in the amino acids of the MAT1-1-1 protein sequences translated from the metatranscriptome assemblies GAGW01008880 and OSIN7648 of natural *C. sinensis* indicated altered overall hydrophobicity-hydrophilicity of the proteins [Bushley *et al*. 2013; Xiang *et al*. 2014; Xia *et al*. 2017].

Table 4. Percentage of amino acids in the MAT1-1-1 and MAT1-2-1 proteins of *H. sinensis* and natural *C. sinensis*.

		% Amino acids of mating-type protein			
		Hydrophobic	**Acidic**	**Basic**	**Neutral**
		MAT1-1-1 protein			
AGW27560	*H. sinensis* Strain CS68-2-1229	41.4%	13.2%	11.0%	34.4%
GAGW01008880	Natural *C. sinensis*	41.3%	14.1%	9.8%	34.8%
OSIN7648	Natural *C. sinensis*	41.9%	13.1%	11.1%	33.7%
		MAT1-2-1 protein			
AFX66389	*H. sinensis* Strain GS09_121	41.2%	10.0%	16.5%	33.3%
OSIN7649	Natural *C. sinensis*	41.2%	10.0%	16.5%	33.3%

Figure 5 shows ExPASy ProtScale plots for the α-Helix, β-Sheet, β-Turn and Coil of the following MAT1-1-1 proteins: the authentic protein AGW27560 of *H. sinensis* Strain CS68-2-1229 and the middle and N-terminal-truncated proteins (OSIN7648 and GAGW01008880, respectively) of natural *C. sinensis*. The variable segment region of the MAT1-1-1 protein is outlined with open boxes in green in all panels of Figure 5. Substantial waveform changes are present in all OSIN7648 panels. The N-terminal-truncated region of the MAT1-1-1 protein GAGW01008880 is marked with open boxes in blue in the AGW27560 and OSIN7648 panels. The alterations in hydrophobicity (Table 4) and the α-Helix, β-Sheet, β-Turn and Coil plots (Figure 5) of the MAT1-1-1 proteins indicate variable folding and secondary structures, indicating disturbed functionality in the fungal mating process and, in turn, the sexual reproduction of *O. sinensis*.


MAT1-1-1 ProtScale Alpha-helix
MAT1-1-1 ProtScale Beta-sheet
alpha-helix / Deleage & Roux
beta-sheet / Chou & Fasman
Score
Position
AGW27560
GAGW01008880
OSIN7648
0.75
0.8
0.85
0.9
0.95
1
1.05
1.1
1.15
1.2
50
100
150
200
250
300
350

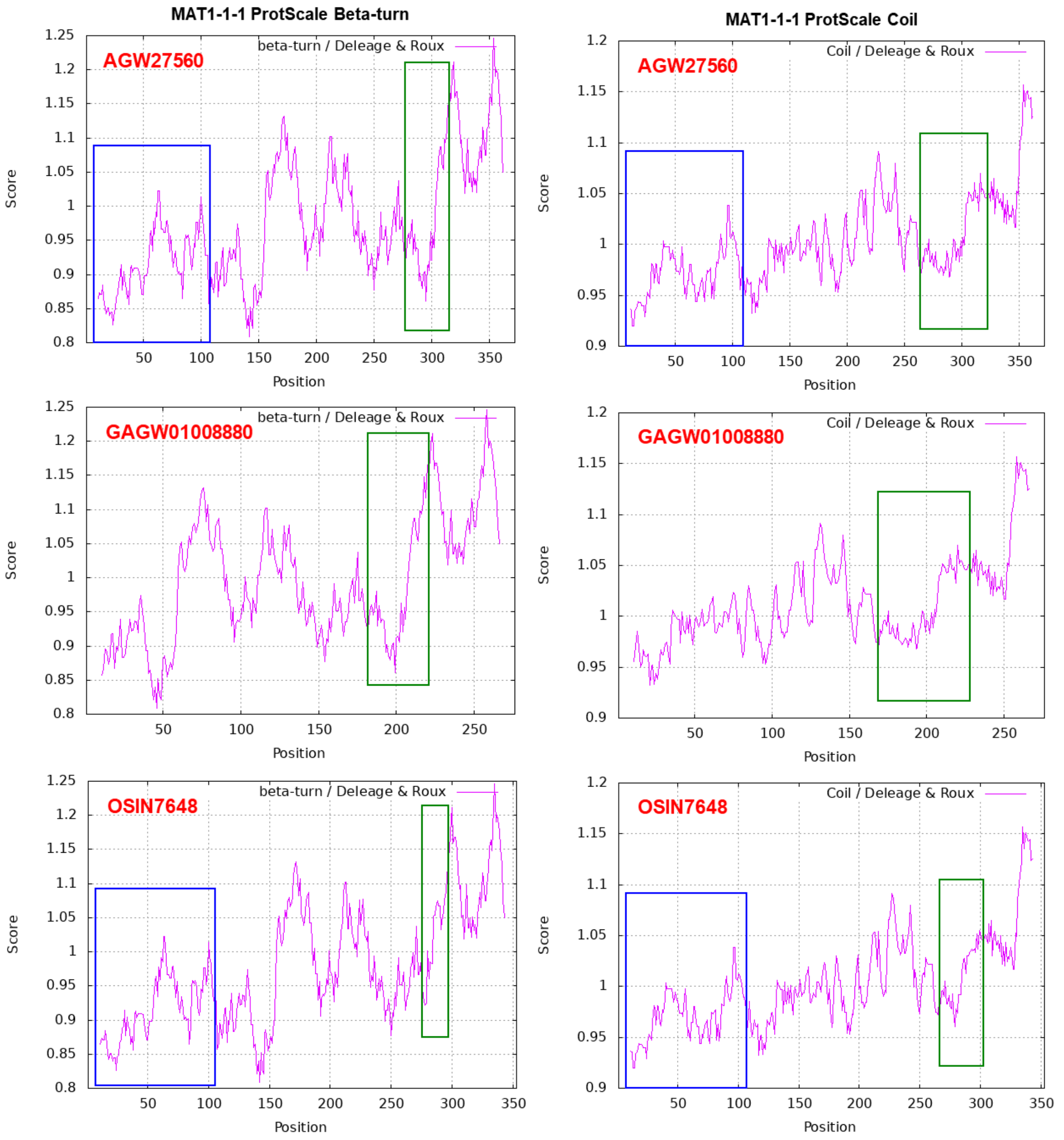


Figure 5. ExPASy ProtScale plots for the MAT1-1-1 protein α-Helix (upper left panels), β-Sheet (upper right panels), β-Turn (lower left panels) and Coil (lower right panels). The MAT1-1-1 protein sequence AGW27560 (372 aa) of *H. sinensis* Strain CS68-2-1229 was compared with the metatranscriptome MAT1-1-1 sequences OSIN7648 (353 aa) and GAGW01008880 (276 aa) of natural *C. sinensis*. Open boxes in green indicate the variable segment regions of the MAT1-1-1 proteins, and open boxes in blue indicate the largely truncated N-terminal segment of the MAT1-1-1 protein GAGW01008880.

The MAT1-2-1 protein sequences derived from 183 *H. sinensis* strains show 98.0–100% homology; 64 strains contain one to five conserved and/or nonconserved mutated residues out of 249 aa residues (Figures 2 and 6) [Zhang *et al*. 2009, 2011, 2014; Bushley *et al*. 2013; Hu *et al*. 2013; Zhang & Zhang 2015]. The MAT1-2-1 protein sequence translated from the metatranscriptome of OSIN7649 obtained from natural *C. sinensis* collected from Deqin County in Yunnan Province shows 100% identity to the MAT1-2-1 protein sequence AFX66389 derived from the genome sequence KC429550 of *H. sinensis* Strain GS09_121 [Zhang *et al*. 2014; Xia *et al*. 2017]. However, the MAT1-2-1 transcript was absent in the metatranscriptome assembly GAGW00000000 of natural *C. sinensis* collected from Kangding County in Sichuan Province (*cf*. Table 3) [Xiang *et al*. 2014]. Table 4 shows no change in the component amino acids of the MAT1-2-1 proteins, AFX66389 of *H. sinensis* Strain GS09_121 and

OSIN7649 of natural *C. sinensis*, indicating unchanged hydrophobicity of the proteins [Zhang *et al*. 2014; Xia *et al*. 2017].

```
AFX66389 1   MANPINMIPNPQWNATDYEAIWKGLEAQVNPFSQILCLEGDFFRQLDDAAKLFIARKLMEHVQESVLYVNDGNGPDRVYLGAPRHFVVGG 90
OSIN7649 1   ------------------------------------------------------------------------------------------ 90
AFX66415 1   ------------------------------------------------------------------------------------------ 90
ACV60375 1   ------------------------------------------------------------------------------------------ 90
ACV60372 1   ---------------------------------L-------------------------------------------------------- 90
AIV43040 1   ---------------I-------------------------------------------------------------------------- 90
AFX66472 1   -----------------------------------------------G------------------------------------------ 90
AFX66473 1   -----------------------------------------------G------------------------------------------ 90
AFX66474 1   -----------------------------------------------G------------------------------------------ 90
AFH35020 1   ---------------I-------------------------------------------------------------------------- 90
ACV60417 1   ---------------I-------------------------------------------------------------------------- 90
ACV60418 1   ---------------I-------------------------------------------------------------------------- 90
AFX66443 1   ---------------I-------------------------------------------------------------------------- 90

AFX66389 91  GMILQISGYAPYWIRRSVSKVVTATVLAPPSPKDIKIPRPPNAYILYRKERHHHVKDANPGITNNEISQILGKAWNMESNDVRQKYKDMS 180
OSIN7649 91  ------------------------------------------------------------------------------------------ 180
AFX66415 91  ------------------------------------------------------------------------------------------ 180
ACV60375 91  ----------------------------------------------------------------------------I------------- 180
ACV60372 91  -----------------------------------------------------Y------------------------------------ 180
AIV43040 91  ----------------------------------------------------------------------------I------------- 180
AFX66472 91  -----------------A------------------------------------------------------------------------ 180
AFX66473 91  -----------------A------------------------------------------------------------------------ 180
AFX66474 91  -----------------A------------------------------------------------------------------------ 180
AFH35020 91  ----------------------------------------------------------------------------I------------- 180
ACV60417 91  ----------------------------------------------------------------------------I------------- 180
ACV60418 91  ----------------------------------------------------------------------------I------------- 180
AFX66443 91  ----------------------------------------------------------------------------I------------- 180

AFX66389 181 QQVKQALLEKHPDYQYKPRRPCERRRRRRASPNQNPKQSTSRNAATRDAAISSEDTSTATGDTNTANGF                      249
OSIN7649 181 ---------------------------------------------------------------------                      249
AFX66415 181 ---------------------------------------------------------------------                      249
ACV60375 181 ---------------------------------------G-----------------------------                      249
ACV60372 181 ---------------------------------------------------------------------                      249
AIV43040 181 -R--------------------------------------------------T-----G----------                      249
AFX66472 181 -----------------------------T-----------------N----T----------------                      249
AFX66473 181 -----------------------------T-----------------N----T----------------                      249
AFX66474 181 -----------------------------T-----------------N----T----------------                      249
AFH35020 181 ---------------------------------------A------------T-----G----------                      249
ACV60417 181 ---------------------------------------A------------T-----G----------                      249
ACV60418 181 ---------------------------------------A------------T-----G----------                      249
AFX66443 181 ---------------------------------------A------------T-----G----------                      249
```

Figure 6. Alignment of the MAT1-2-1 protein sequences translated from transcript sequences of *H. sinensis* strains and the metatranscriptome sequence of natural *C. sinensis*. OSIN7649 is the MAT1-2-1 protein sequence derived from the metatranscriptome sequence of natural *C. sinensis* collected from Deqin County of Yunnan Province, China [Xia *et al*. 2017]. AFX66389, AFX66415, ACV60375, ACV60372, AIV43040, AFX66472, AFX66473, AFX66474, AFH35020, ACV60417, ACV60418 and AFX66443 are the MAT1-2-1 protein sequences of *H. sinensis* Strains GS09_121, QH09_201, XZ-SN-44, XZ-LZ06-61, XZ12_16, YN09_6, YN09_22, YN09_51, XZ06-124, XZ-LZ07-H1, XZ-LZ07-H2 and XZ05_8, respectively [Zhang *et al*. 2009, 2011, 2014; Bushley *et al*. 2013]. The residues in red indicate the variant amino acids, and hyphens indicate identical amino acid residues.

In addition to the above protein sequence analyses, Zhang and Zhang [2015] reported 4.7% and 5.7% allelic variations in the coding sequences of the MAT1-1-1 and MAT1-2-1 genes, respectively, resulting in 5.9% and 5.6% variations in the MAT1-1-1 and MAT1-2-1 protein sequences, respectively. The allelic and amino acid variations are much greater than what we observed through GenBank database analysis, suggesting that Zhang and Zhang [2015] may not have uploaded all gene and protein sequences that they analyzed to GenBank. It is conceivable that the mutant MAT1-1-1 and MAT1-2-1 proteins of natural *C. sinensis* might not necessarily be produced by Genotype #1 *H. sinensis* of *O. sinensis* but might be produced by one of multiple genotypes of *O. sinensis* fungi and cocolonized fungi when considering the failure to detect mating-type genes and transcripts within the *MAT1-1* and *MAT1-2* idiomorphs in numerous *H. sinensis* strains or wild-type isolates from natural *C. sinensis* (*cf*. Tables 1 and 3) and "to induce development of the *C. sinensis* fruiting bodies" during 40 years of experience using pure *H. sinensis* cultures in research-oriented academic settings, as summarized by Zhang *et al*. [2013].

Occurrence and Transcription of Pheromone Receptor Genes and Variations in Pheromone Receptor Proteins in H. sinensis *and Natural* C. sinensis

Heterothallic sexual development in morphologically indistinguishable fungal haploid cells is controlled by the growth of complementary mating mycelia that produce complementary mating proteins followed by gametangial contact and the fusion of hyphal or ascosporic cells to induce plasmogamy and karyogamy in fungi [Jones & Bennett 2011; Zheng & Wang 2013; Wilson *et al*. 2015]. These mating processes for ascomycetes occur between two types of fungal cells, a and α cells, and the mating of these fungal cells is regulated by secreted pheromones. Pheromones play essential roles in choreographing the interactions between mating partners by regulating intercellular mating communication and conveying information from one type of cell to another cell type of the same species for physiological heterothallism or from a different species for hybridization. However, the biochemical processes for a- and α-pheromone precursors in ascomycetes appear to be complex and are not fully understood for *O. sinensis*. To date, scientists have not identified the genes encoding a- and α-pheromones in the genome assemblies of *H. sinensis* Strains 1229, CC1406-203, Co18, IOZ07 and ZJB12195 [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020].

An α-pheromone is recognized by a-type pheromone receptors (class 4 G-protein-coupled receptors) on the surface of an a-cell and *vice versa* [Jones & Bennett 2011; Zheng & Wang 2013; Wilson *et al*. 2015]. The activated receptor plays crucial roles in the reciprocal interaction of mate recognition and the initiation of signaling and mating processes.

a-Factor-Like Pheromone Receptor in *H. sinensis* and Natural *C. sinensis*

Hu *et al*. [2013] reported the occurrence of an a-pheromone receptor gene (*PRE-1*; 3→1,386 of KE659055) in *H. sinensis* Strain Co18. It contains 3 exons separated by introns I and II (437→490 and 1,018→1,075 of KE659055, respectively) (Figure 7). The *PRE-1* gene sequences of the genome assemblies LKHE01002330 (1,592→2,975), NGJJ01000446 (84,292←85,675), ANOV01020521 (3→1,386) and JAAVMX010000007 (3,515,715→3,517,197) of *H. sinensis* Strains 1229, CC1406-203, Co18 and IOZ07, respectively, show 100% homology with each other [Hu *et al*. 2013; Li *et al*. 2016a; Liu *et al*. 2020; Shu *et al*. 2020]. However, the sequence of the *PRE-1* gene exhibited 95.3% similarity with the genome assembly LWBQ01000004 (451,200→452,640) of *H. sinensis* Strain ZJB12195 due to multiple insertions and several transversion alleles in exon I (Table 5, Figure 7) [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020]. The coding sequence of *PRE-1* is absent in the transcriptome assembly GCQL00000000 of *H. sinensis* Strain L0106 [Liu *et al*. 2015], indicating either silent transcription or the complete absence of the gene in a-cells in *H. sinensis* Strain L0106, resulting in impaired mate recognition upon interaction with the α-factor pheromone.

Table 5. Percentage of similarity between the protein sequences of pheromone receptors and the genome and transcriptome sequences of *H. sinensis* and natural *C. sinensis*.

	Percent similarity	
	a-Factor-like pheromone receptor	**α-Factor-like pheromone receptor**
	Between the sequences of pheromone receptor genes and the *H. sinensis* genomes	
H. sinensis Strain 1229	100%	99.8%
H. sinensis Strain CC1406-203	100%	99.8%
H. sinensis Strain Co18	100%	100%
H. sinensis Strain IOZ07	100%	99.9%
H. sinensis Strain ZJB12195	95.2%	97.5%
	Between the coding sequences of the pheromone receptor genes (excluding introns) and the transcriptome sequences of *H. sinensis* and natural *C. sinensis*	
H. sinensis Strain L0106	—	99.3%
Natural *C. sinensis* (Deqin, Yunnan)	100%	99.7%
Natural *C. sinensis* (Kangding, Sichuan)	98.5–100%	—
	Between the pheromone receptor protein sequences translated from the gene sequences and the transcript sequences of *H. sinensis* and natural *C.* sinensis	
H. sinensis Strain L0106	—	86.4–100%

Natural *C. sinensis* (Deqin, Yunnan)	100%	82.9%
Natural *C. sinensis* (Kangding, Sichuan)	97.8–100%	—

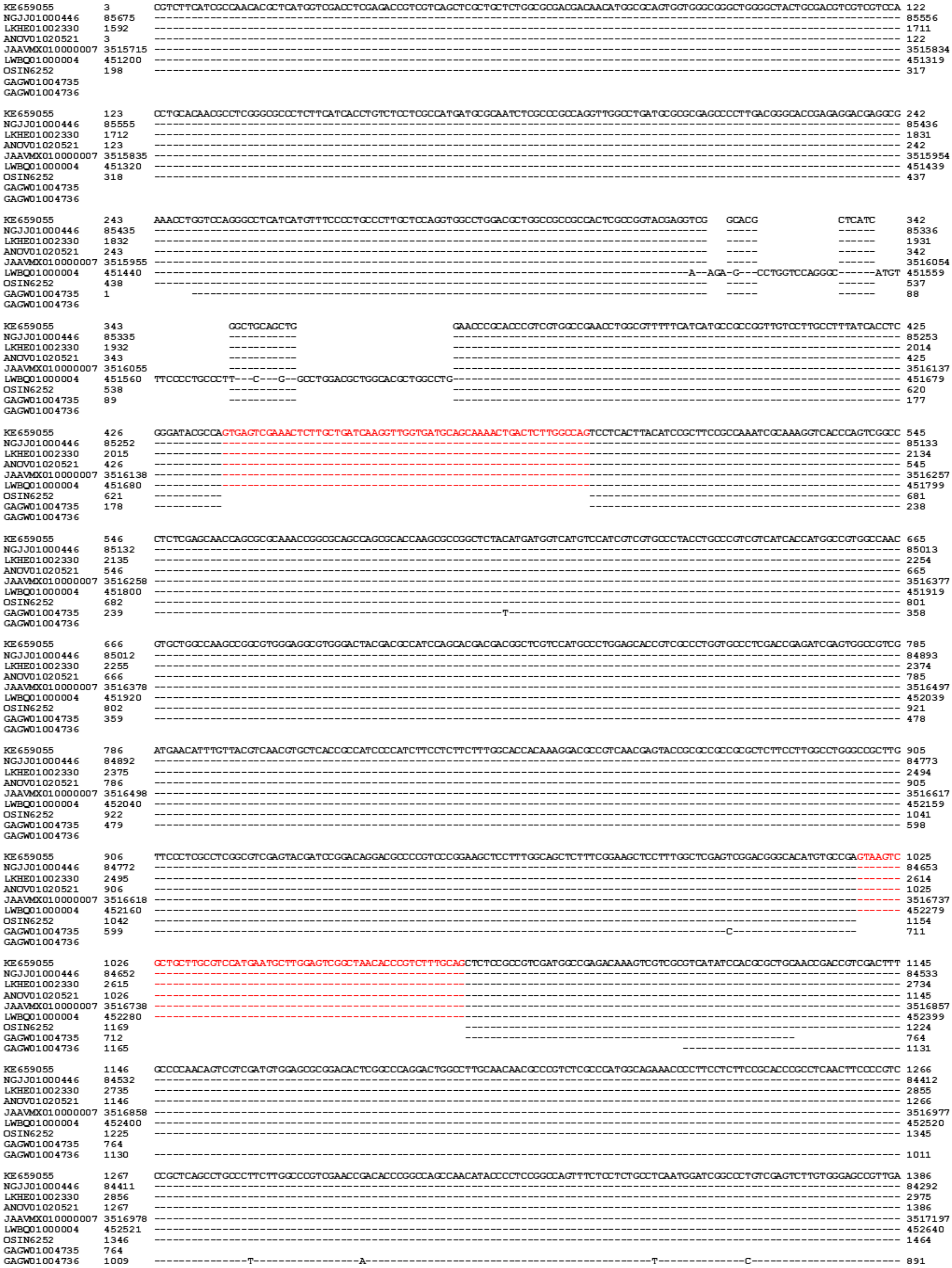


Figure 7. Alignment of the genome and transcriptome sequences of a-pheromone receptor (*PRE-1*) from *H. sinensis* strains and natural *C. sinensis*. KE659055 is the gene sequence of *H. sinensis* Strains Co18 [Hu *et al*. 2013]. The genome assemblies NGJJ01000446, LKHE01002330, ANOV01020521, JAAVMX010000007, and LWBQ01000004 were obtained from *H. sinensis* strains CC1406-203, 1229, Co18, IOZ07, and ZJB12195, respectively [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Jiu *et al*. 2020; Shu *et al*. 2020]. OSIN6252 and GAGW01004735/GAGW01004736 are the transcriptome assembly

sequences obtained from natural *C. sinensis* specimens collected from Deqin County of Yunnan Province and Kangding County of Sichuan Province of China, respectively [Xiang *et al*. 2014; Xia *et al*. 2017]. The segments in red are introns I and II. Hyphens indicate identical bases, and spaces denote unmatched sequence gaps.

The *PRE-1* gene sequence is present in the metatranscriptome assemblies OSIN6252 and GAGW01004735/GAGW01004736 of natural *C. sinensis* specimens collected from Deqin County of Yunnan Province and Kangding County of Sichuan Province in China, respectively, showing 98.5–100% homology to the coding sequence of the *PRE-1* gene KE659055 of *H. sinensis* Strain Co18 (Table 5) [Hu *et al*. 2013; Xiang *et al*. 2014; Xia *et al*. 2017].

The metatranscript OSIN6252 (220→1461) is 100% identical to the a-pheromone receptor gene sequence KE659055 of *H. sinensis* Strain Co18 (Figure 7). The metatranscript GAGW01004735 (1→764) is missing a 232-bp segment in exon I at its 5′ end and a 257-bp segment in exon III at its 3′ end, while the metatranscript GAGW01004736 (1029←1165) covers the missing portion of exon III at the 3′ end of the a-pheromone receptor gene sequence (Figure 7). Although the fragmentation of GAGW01004735 and GAGW01004736 may represent an assembly error if they are assumed to be derived from the same fungal genome [Xiang *et al*. 2015], the integration of the two fragments is still 232 bp short at the 5′ end of exon I of the *PRE-1* gene.

In addition to the complete absence of the *PRE-1* transcript and protein in *H. sinensis* Strain L0106 [Liu *et al*. 2015], the protein sequences translated from the metatranscriptome assembly OSIN6252 of natural *C. sinensis* are 100% identical to the protein sequence EQK97482 of the a-pheromone receptor that was directly translated from the *PRE-1* gene KE659055 of *H. sinensis* Strain Co18 (Table 5; Figures 7 and 8) [Hu *et al*. 2013; Xia *et al*. 2017]. The metatranscriptome GAGW01004735/GAGW01004736 of natural *C. sinensis* is 99.4% homologous to the protein sequence EQK97482 of the a-pheromone receptor (Table 5) [Hu *et al*. 2013; Xiang *et al*. 2014]. The translated *PRE-1* protein sequence of GAGW01004735 is missing 75 aa and 85 aa at its N- and C-termini, respectively, whereas the other metatranscript, GAGW01004736, covers the C-terminal portion of the *PRE-1* protein with two nonconserved mutated aa residues (Figure 8).

Table 6 lists the 4 types of amino acids in α-pheromone receptor proteins based on their chemical-physical and hydrophobicity properties using Peptide 2.0 [Peters & Elofsson 2014; Simm *et al*. 2016]. The segment deletion and nonconserved aa mutations in the *PRE-1* protein sequence produce a truncated a-pheromone receptor protein without significant changes in hydrophobicity (50.4–50.7% hydrophobic amino acids) (Table 6) or in α-Helix, β-Sheet, β-Turn and Coil plots of the a-pheromone receptor proteins (GAGW01004735-GAGW01004736 and OSIN6252) in natural *C. sinensis* (data not shown), which raises the question of whether the mutations cause altered conformation of the membrane proteins and their functionality in interindividual mating signal communication between mating partners.

```
EQK97482      1     MVDLETVVSSLLWRDDNMAQWWAGWGYCDVVVHLHNASGALFITCLLAMMRNLARQVGLM  60
KAF4506800    74    ------------------------------------------------------------  133
OSIN6252      220   ------------------------------------------------------------  399
GAGW01004735
GAGW01004736

EQK97482      61    RASPLTGTERTRRNLVQGLIMFPLPLLQVAWTLAAATRRYEVGTLIGCSWNPHPSWPNLA  120
KAF4506800    134   ------------------------------------------------------------  193
OSIN6252      400   ------------------------------------------------------------  579
GAGW01004735  2                  -----------------------------------------------  136
GAGW01004736

EQK97482      121   FFIMPPVVLAFITSGYAILTYIRFRQIAKVTQSALSSNQRANRRSQRTKRRLYMMVMSIV  180
KAF4506800    194   ------------------------------------------------------------  253
OSIN6252      580   ------------------------------------------------------------  759
GAGW01004735  137   ------------------------------------------------------------  316
GAGW01004736

EQK97482      181   VPYLPVVITMAVANVLAKPAWEAWDYDAIQHDDGSSMPWSTVALVPSTEIEWPSMNICYV  240
KAF4506800    254   ------------------------------------------------------------  313
OSIN6252      760   ------------------------------------------------------------  939
GAGW01004735  317   ------------------------------------------------------------  496
GAGW01004736

EQK97482      241   NVLTAIPIFLFFGTTKDAVNEYRAAALFLGLGRLFPRLGVEYDPDRTPRPGSSFGSSFGS  300
KAF4506800    314   ------------------------------------------------------------  373
OSIN6252      940   ------------------------------------------------------------  1119
GAGW01004735  497   ------------------------------------------------------------  676
GAGW01004736

EQK97482      301   SFGSSRTGTCADSPPSMAETKSSRHIHALQPTVDFAPTVVDVERGHSAQDWPCNNARLAH  360
KAF4506800    374   ------------------------------------------------------------  433
OSIN6252      1120  ------------------------------------------------------------  1299
GAGW01004735  677   ------------------------------                                763
GAGW01004736  1163                       ---------------------------------------  1128

EQK97482      361   GRNPFLFRTRLNFPVPLSLPFLARRTDTRPANIPLRPVSPLPQWIGPVESCGSR        360
KAF4506800    434   ------------------------------------------------------        433
OSIN6252      1300  ------------------------------------------------------        1461
GAGW01004735
GAGW01004736  1163  --------------S-----T---------------------------------        1128
```

Figure 8. Alignment of the translated protein sequences of a-pheromone receptor (*PRE-1*) from *H. sinensis* strains and natural *C. sinensis*. The *PRE-1* protein sequences EQK97482 and KAF4506800 are translated from the gene sequence KE659055 of *H. sinensis* strain Co18 and the genome segment sequence JAAVMX010000007 of *H. sinensis* strain IOZ07 [Hu *et al*. 2013; Shu *et al*. 2020]. The *PRE-1* transcriptome assembly sequences OSIN6252, GAGW01004735 and GAGW01004736 of natural *C. sinensis* are translated into protein sequences [Xiang *et al*. 2014; Liu *et al*. 2015]. Hyphens indicate identical bases, and spaces denote unmatched sequence gaps.

Table 6. Percentages of amino acids of the pheromone receptor proteins of *H. sinensis* and natural *C. sinensis*.

		% Amino acids of pheromone receptor protein			
		Hydrophobic	**Acidic**	**Basic**	**Neutral**
		a-pheromone receptor protein			
EQK97482	*H. sinensis* Strain Co18	50.7%	6.5%	11.6%	31.4%
GAGW01004735-GAGW01004736	Natural *C. sinensis*	50.4%	6.2%	11.2%	31.2%
OSIN6252	Natural *C. sinensis*	50.7%	6.5%	11.6%	31.4%
		α-pheromone receptor protein			
EQK99119	*H. sinensis* Strain Co18	55.1%	5.6%	7.4%	31.8%
GCQL01017756	*H. sinensis* Strain L0106	50.8%	6.0%	9.0%	34.2%
GCQL01015779	*H. sinensis* Strain L0106	51.8%	6.0%	9.0%	33.2%
OSIN6424	Natural *C. sinensis*	52.7%	5.8%	8.6%	32.8%

α-Factor-like Pheromone Receptor in *H. sinensis* and Natural *C. sinensis*

Hu *et al*. [2013] reported the occurrence of the α-pheromone receptor gene (2,275→3,724 of KE653642) in *H. sinensis* Strain Co18. This receptor gene exhibits 97.5–100% homology with the genome assemblies LKHE01001069 (13,389→13,603)/LKHE01000580 (23,846←24,749; 23,657←23,794), NGJJ01001310 (128,759←130,010), ANOV01006352 (1→138)/ANOV01006351 (2,275→3,528), JAAVMX010000012 (406,964→408,405) and LWBQ01000044 (488,856→490,264) in *H. sinensis* Strains 1229, CC1406-203, Co18, IOZ07 and ZJB12195, respectively (Figure 9, Table 4) [Hu *et* *al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020]. Introns I and II of the α-pheromone receptor gene are located at sites 3,091→3,240 and 3,528→3,586 of the sequence KE653642 of Strain Co18 [Hu *et al*. 2013]. A 133-bp segment deletion in exon I of the gene occurs between the genome assembly segments of LKHE01001069 and LKHE01000580 of *H. sinensis* Strain 1229 (Figure 9), which encodes a 45-aa deleted segment (ALVVAFAVLALTPAAKLRRPSSLLHLAGLAMCLARVGSLAVPA-LS) in the protein sequence.

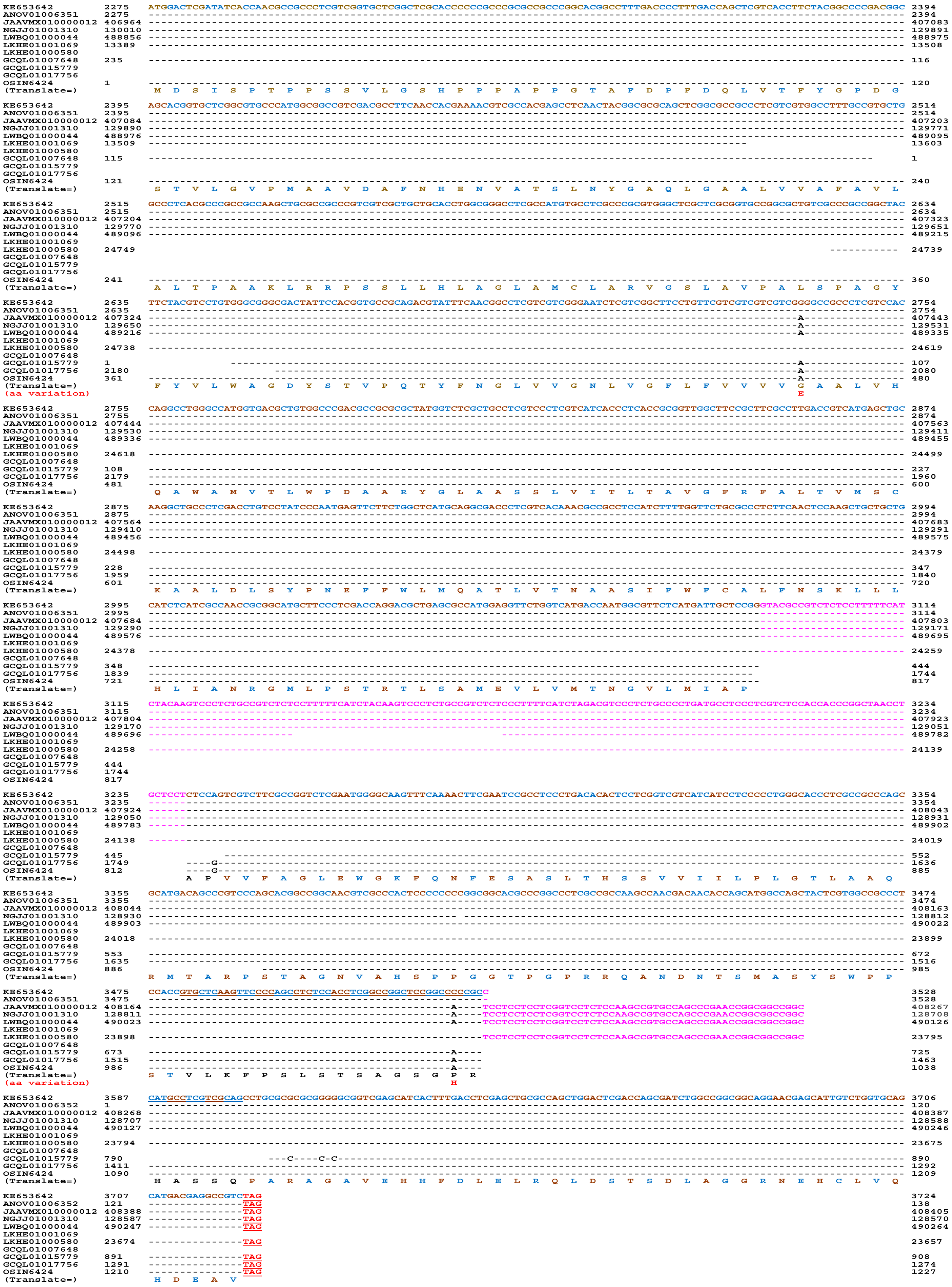


Figure 9. Alignment of the genome and transcriptome sequences of the α-pheromone receptor of *H. sinensis* and natural *C. sinensis*. The genome sequence of KE653642 was obtained from *H. sinensis* Strain Co18 [Hu *et al*. 2013]. JAAVMX010000012, NGJJ01001310, LWBQ0100044, ANOV01006352, and LKHE01000580 are genome assemblies obtained from *H. sinensis* strains IOZ07, CC1406-203, ZJB12195, Co18, and 1229, respectively [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Jiu *et al*. 2020; Shu *et al*. 2020]. GCQL01007648, GCQL01017756, GCQL01015779, and OSIN6424 are the transcriptome assemblies obtained from *H. sinensis* strain L0106 and natural *C. sinensis* specimens collected from Deqin County of Yunnan Province, China [Liu *et al*. 2015; Xia *et al*. 2017]. The triplets shown in alternating brown and blue indicate the open reading frame. The sequences in pink indicate introns I and II between nucleotides 3,092→3,240 and 3,528→3,586, respectively, of the α-pheromone receptor gene KE653642, and the underlined "TAG" triplets shown in red are stop codons. The protein sequence was translated from the transcript sequences. The amino acid residues in red represent variants. Hyphens indicate identical bases, and spaces denote unmatched sequence gaps.

The α-pheromone receptor gene contains 3 exons and 2 introns (Figure 9). The coding sequence of the gene of *H. sinensis* Strain Co18 shows 99.3–99.7% homology to the transcriptome assemblies GCQL01007648/GCQL01017756/GCQL01015779 of *H. sinensis* Strain L0106 and the metatranscriptome assembly OSIN6424 of natural *C. sinensis* collected from Deqin County of Yunnan Province, China (Table 4) [Hu *et al*. 2013; Liu *et al*. 2015; Xia *et al*. 2017]. GCQL01007648 (1←235) covers only 235 bp at the 5′ end of the 720 bp exon I. GCQL01015779 (1→908) and GCQL01017756 (1,274←2,180) cover 444 bp and 436 bp, respectively, at the 3′ end of exon I, and the full-length sequences of exons II and III. There are two gaps (138 bp and 145 bp) between the sequences of GCQL01007648 and GCQL01017756 and between those of GCQL01007648 and GCQL01015779 [Liu *et al*. 2015].

The presence of the α-pheromone receptor transcript but not the a-pheromone receptor transcript in *H. sinensis* Strain L0106 suggests that the strain can only receive a mating signal from a sexual partner that contains a-cells and produces a-pheromone. The differential transcription of pheromone receptor genes in *H. sinensis* Strain L0106 (Table 4), together with the differential transcription of mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs (*cf*. Table 3), indicates that *H. sinensis* Strain L0106 needs a sexual partner that produces a-pheromone and protein(s) of the *MAT1-1* idiomorph for sexual reproduction. The sexual partner may be the same species for physiological heterothallism or a different species for hybridization when homothallic or pseudohomothallic reproduction is impossible, as described above.

The metatranscriptome assembly OSIN6424 of natural *C. sinensis* contains the sequences of all 3 exons of the α-pheromone receptor gene with a 58-bp deletion (GTCGTCATCATCCTCCCCCTGGGCA-CCCTCGCCGCCCAGCGCATGACAGCCCGTCCCA) between nucleotides 885 and 886 in exon II of the cDNA (*cf*. Figure 9) [Xia *et al*. 2017]. Notably, the deletion of 58 bp is not a multiple of 3 and may represent a frameshift mutation that involves a peptide segment of 52 aa encoded by the 3′ end region of exon II with no stop codon, while intron II of the gene is likely in "Intron Phase 0", and the frameshift mutation may not involve exon III of the gene.

The α-pheromone receptor gene is absent in the metatranscriptome assembly GAGW00000000 of natural *C. sinensis* collected from Kangding County of Sichuan Province [Xiang *et al*. 2014].

The translated protein sequence of the metatranscriptome assembly OSIN6424 (1→1,224) from natural *C. sinensis* shows 82.9% similarity to the α-pheromone receptor protein sequence EQK99119, which was translated directly from the segment sequence of KE653642 of *H. sinensis* Strain Co18. OSIN6424 exhibits an insertion (38 aa), two segment deletions (9 and 11 aa) and 15-aa residues of nonconserved mutations (Figure 10, Table 4) [Hu *et al*. 2013; Xia *et al*. 2017].

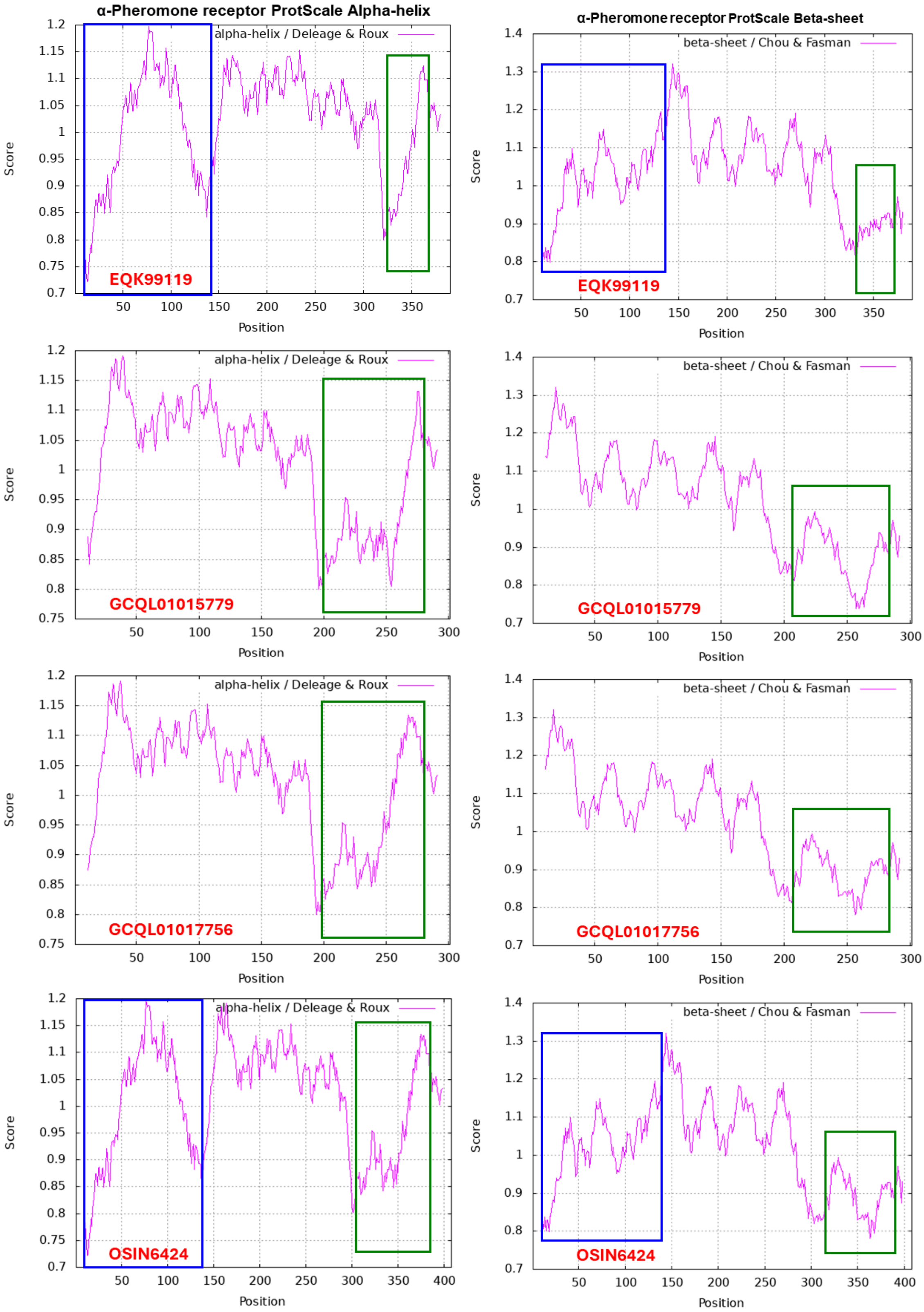
```
EQK99119     1    MDSISPTPPSSVLGSHPPPAPPGTAFDPFDQLVTFYGPDGSTVLGVPMAAVDAFNHENVA 60
OSIN6424     1    ------------------------------------------------------------ 180
GCQL01007648 235  ------------------------------------------------------------ 56
GCQL01017756
GCQL01015779

EQK99119     61   TSLNYGAQLGAALVVAFAVLALTPAAKLRRPSSLLHLAGLAMCLARVGSLAVPALSPAGY 120
OSIN6424     181  ------------------------------------------------------------ 360
GCQL01007648 55   ------                                                       38
GCQL01017756
GCQL01015779

EQK99119     121  FYVLWAGDYSTVPQTYFNGLVVGNLVGFLFVVVVGAALVHQAWAMVTLWPDAARYGLAAS 180
OSIN6424     361  ----------------------------------E------------------------- 540
GCQL01007648 38
GCQL01017756 2179        ---------------------------E------------------------- 2021
GCQL01015779 3         -----------------------------E------------------------- 167

EQK99119     181  SLVITLTAVGFRFALTVMSCKAALDLSYPNEFFWLMQATLVTNAASIFWFCALFNSKLLL 240
OSIN6424     541  ------------------------------------------------------------ 720
GCQL01007648 38
GCQL01017756 2020 ------------------------------------------------------------ 1841
GCQL01015779 168  ------------------------------------------------------------ 347

EQK99119     241  HLIANRGMLPSTRTLSAMEVLVMTNGVLMIAPVVFAGLEWGKFQNFESASLTHSSVVIIL 300
OSIN6424     721  -------------------------------------------------------      885
GCQL01007648 38
GCQL01017756 1840 ------------------------------------------------------------ 1661
GCQL01015779 348  ------------------------------------------------------------ 527

EQK99119     301  PLGTLAAQRMTARPSTAGNVAHSPPGGTPGPRRQANDNTSMASYSWPPST           350
OSIN6424     886      HGR--RPLP-GGTPGPR           ------------------VLKFPSLSTS 1020
GCQL01007648 38
GCQL01017756 1660 --------------------------------------------------VLKFPSLSTS 1481
GCQL01015779 528  --------------------------------------------------VLKFPSLSTS 707

EQK99119     351                              PARAGAVEHHFDLELRQLDSTSDLAGGRNEHC 382
OSIN6424     1021 AGSGHRSSSSVLSKPCQPEPAAGHASSQ-------------------------------- 1200
GCQL01007648 38
GCQL01017756 1480 AGSGHRSSSSVLSKPCQPEPAAGHASSQ-------------------------------- 1301
GCQL01015779 708  AGSGHRSSSSVLSKPCQPEPAGRPPPPP----  -------------------------- 881

EQK99119     383  LVQHDEAV                                                     390
OSIN6424     1201 --------                                                     1224
GCQL01007648 38
GCQL01017756 1300 --------                                                     1277
GCQL01015779 882  --------                                                     905
```

Figure 10. Alignment of the protein sequences of the α-pheromone receptor of *H. sinensis* strain Co18 and the transcriptome sequences of *H. sinensis* strain L0106 and natural *C. sinensis*. The gene sequence EQK99119 was obtained from *H. sinensis* strains Co18 [Hu *et al*. 2013]. The transcriptome assemblies GCQL01007648, GCQL01017756, GCQL01015779 and OSIN6424 were obtained from *H. sinensis* strain L0106 and natural *C. sinensis* specimens collected from Deqin County of Yunnan Province, China [Xiang *et al*. 2014; Xia *et al*. 2017]. Hyphens indicate identical amino acid residues, and spaces denote unmatched sequence gaps.

The translated protein sequence of GCQL01007648 (38←235) of *H. sinensis* Strain L0106 is 100% identical to the small N-terminal portion (66 aa) of EQK99119 [Hu *et al*. 2013; Liu *et al*. 2015]. The translated protein sequences of GCQL01017756 and GCQL01015779 overlap and cover the majority of the protein sequence through the C-terminus. These sequences show 86.4% and 87.0% similarity to EQK99119 with a deletion of the 59–61 aa residues AQLGAALVVAFAVLALTPAAKLRRPSSLLHLAGLAMCLAR-VGSLAVPALSPAGYFYVLW(AG) and several conserved and/or nonconserved mutated residues.

Compared with the percentages of amino acids of the α-pheromone receptor protein EQK99119 of *H. sinensis* Strain Co18 as shown in Table 6, changes in the amino acids of α-pheromone receptor protein sequences translated from the transcriptome assemblies GCQL01017756 and GCQL01015779 of *H. sinensis* Strain L0106 and the metatranscriptome assembly OSIN6424 of natural *C. sinensis* indicated dramatic decreases in hydrophobic amino acids and complementary increases in acidic, basic and neutral amino acids [Hu *et al*. 2013; Liu *et al*. 2015; Xia *et al*. 2017].

Figure 11 shows the ExPASy ProtScale plots for the α-Helix, β-Sheet, β-Turn and Coil of the α-pheromone receptor proteins: the authentic protein EQK99119 of *H. sinensis* Strain Co18 and the variable α-pheromone receptor sequences translated from the transcriptome assemblies GCQL01017756 and GCQL01015779 of *H. sinensis* Strain L0106 and the metatranscriptome assembly OSIN6424 of natural *C. sinensis*. Substantial waveform changes were found in all the GCQL01017756, GCQL01015779 and OSIN6424 panels and are outlined with open boxes in green. The N-terminal-truncated region of the proteins is marked with open boxes in blue in the EQK99119 and OSIN6424 panels.



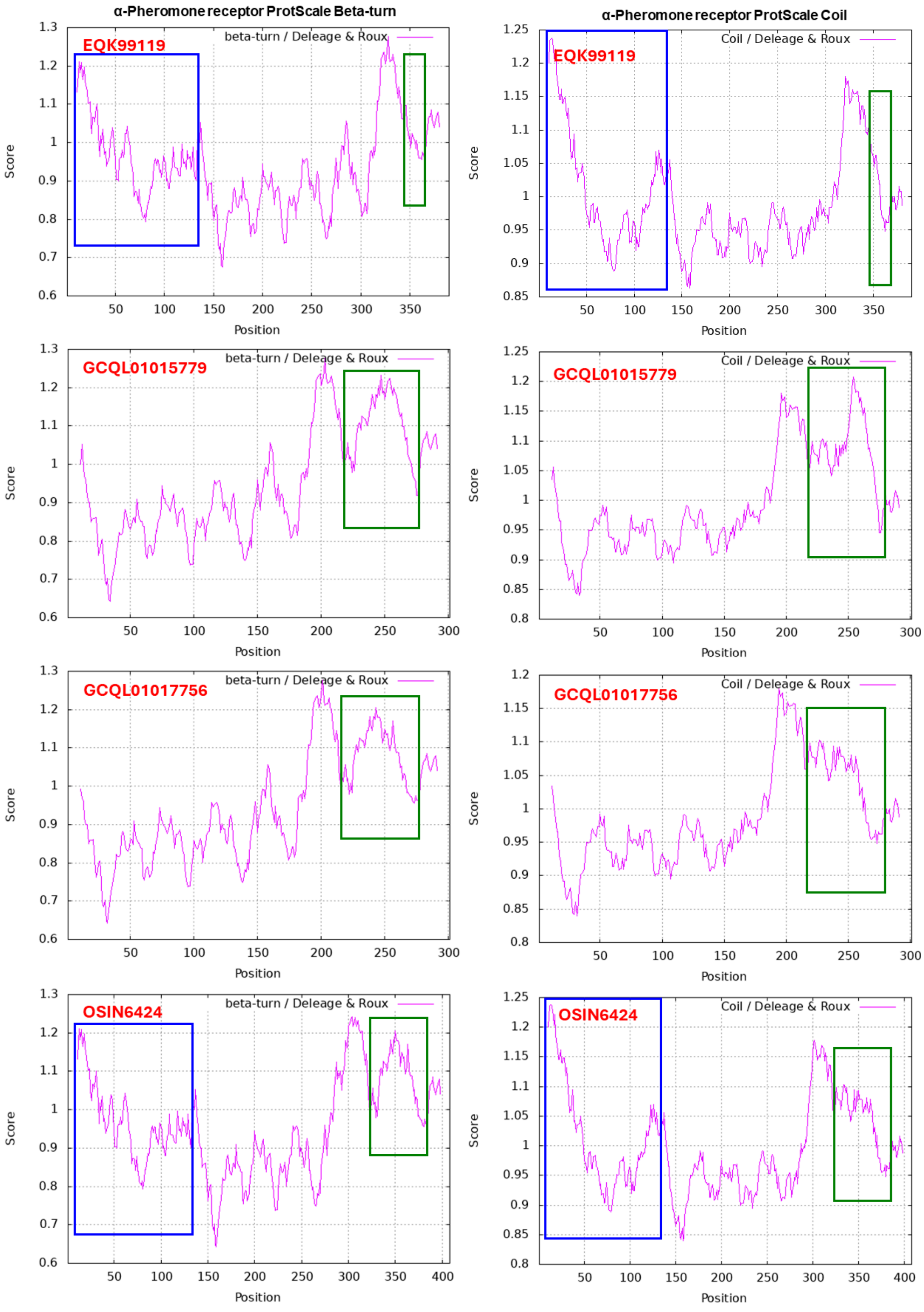


Figure 11. ExPASy ProtScale plots for the α-Helix (upper-left panels), β-Sheet (upper-right panels), β-Turn (lower-left panels) and Coil (lower-right panels) of α-pheromone receptor proteins. The α-pheromone receptor protein sequence EQK99119 (390 aa) of *H. sinensis* Strain Co18 was compared with the transcriptome and metatranscriptome α-pheromone receptor sequences and GCQL01017756 (301 aa) and GCQL01015779 (301 aa) of *H. sinensis* Strain L0106 and OSIN6424 (408 aa) of natural *C. sinensis*. Open boxes in green indicate the variable segment regions of the α-pheromone receptor

proteins, and open boxes in blue indicate the largely truncated N-terminal segment of the α-pheromone receptor proteins GCQL01017756 and GCQL01015779.

Zheng & Wang [2013] reported that the α-pheromone receptor is a G protein-coupled receptor possessing 7 transmembrane domains. The membrane protein of the α-pheromone receptor EQK99119 of *H. sinensis* Strain Co18 contains 55.1% hydrophobic amino acids (*cf*. Table 6), which is significantly greater than the approximately 41% hydrophobic amino acids in MAT1-1-1 proteins (*cf*. Table 4). The mutant α-pheromone receptor proteins in *H. sinensis* Strain L0106 and natural *C. sinensis* contain fewer hydrophobic amino acids (50.8–52.7%). The alterations in the hydrophobicity (*cf*. Table 6) and the α-Helix, β-Sheet, β-Turn and Coil plots (*cf*. Figure 11) of the α-pheromone receptor proteins of *H. sinensis* Strain L0106 and natural *C. sinensis* indicate variable folding and secondary structures of the receptor proteins, indicating altered functionality of the sexual signal reception from mating partners and, in turn, the sexual reproduction of *O. sinensis,* probably favoring hybridization.

Other Pheromone-Related Genes in *H. sinensis* and Natural *C. sinensis*

Hu *et al*. [2013] also identified genes encoding a pheromone-regulated membrane protein (65,257→66,887 of KE652182), the a-pheromone processing metallopeptidase Ste23 (10361→13752 of KE652396), and other related genes in *H. sinensis* Strain Co18. The proteins encoded by these genes participate in the biological processes of mating protein activation and signal transduction. The segment sequences of KE652182 and KE652396 exhibit 99.9–100% homology to the genome assemblies LKHE01001238 (91,967→93,597)/LKHE01000585 (20,198←23,589), NGJJ01000102 (63,761→65,391)/NGJJ01001093 (42,301→45,691), ANOV01000086 (17,505→19,135)/ANOV01001222 (10,361→13,752) and JAAVMX010000009 (1,427,620→1,429,250)/JAAVMX010000008 (1,724,511→1,727,903) of *H. sinensis* Strains 1229, CC1406-203, Co18 and IOZ07, respectively, but 95.8–98.4% similarity to the genome assemblies LWBQ01000047 (64,210→65,902)/LWBQ01000028 (337,499←339,681) of Strain ZJB12195 due to multiple insertion/deletion mutations and other transition and transversion mutations in alleles [Hu *et al*. 2013; Li *et al*. 2016a; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020].

The KE652182 gene encoding a pheromone-regulated membrane protein shows 95.2–96.3% similarity to the transcriptome assembly GCQL01011718 of *H. sinensis* Strain L0106 and the metatranscriptome assemblies GAGW01006658 and OSIN1278 of natural *C. sinensis* collected from Kangding County of Sichuan Province and Deqin County of Yunnan Province, respectively, with five to six deletion gaps [Hu *et al*. 2013; Xiang *et al*. 2014; Liu *et al*. 2015; Xia *et al*. 2017].

The KE652396 gene encoding the a-pheromone processing metallopeptidase Ste23 exhibits 99.2–100% homology to the transcriptome assembly GCQL010123128 of *H. sinensis* Strain L0106 and the metatranscriptome assemblies GAGW01006419/GAGW01005794 and OSIN0988 of natural *C. sinensis* collected from Kangding County of Sichuan Province and Deqin County of Yunnan Province, respectively [Hu *et al*. 2013; Xiang *et al*. 2014; Liu *et al*. 2015; Xia *et al*. 2017].

Discussion

Reproductive Behavior of H. sinensis, Genotype #1 of O. sinensis

Three reproduction hypotheses have been previously proposed for *H. sinensis*, the postulated sole anamorph of *O. sinensis* [Zhu & Li 2017; Li *et al*. 2022a, 2022b]: homothallism [Hu *et al*. 2013], pseudohomothallism [Bushley *et al*. 2013] and facultative hybridization [Zhang & Zhang 2015]. In theory, self-fertilization in ascomycetes becomes a reality when the mating-type genes of both the *MAT1-1* and *MAT1-2* idiomorphs are successfully transcribed and translated, and the mating proteins are synthesized and fully activated within a single fungal cell [Turgeon & Yoder 2000; Debuchy & Turgeo 2006; Jones & Bennett 2011; Zhang *et al*. 2011; Bushley *et al*. 2013; Hu *et al*. 2013; Zheng & Wang 2013; Wilson *et al*. 2015; Zhang & Zhang 2015]. According to this study, the differential occurrence of the mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs in 237 *H. sinensis* strains (*cf*. Table 1) fails to support the genetic-based capability of self-fertilization. The transcriptome assembly GCQL00000000 of *H. sinensis* Strain L0106 contains a transcript of the MAT1-2-1 gene but no transcripts of the *MAT1-1* idiomorph [Liu *et al*. 2015]. The differential occurrence and transcription of mating-type genes at the genomic and transcriptomic levels are not consistent with the self-fertilization hypothesis for *H. sinensis* under homothallic or pseudohomothallic reproduction; instead, they support a reproductive strategy of either physiological heterothallism or hybridization [Turgeon & Yoder 2000; Pfennig 2007; Zheng & Wang 2013; Wilson *et al*. 2015; Du *et al*. 2020; Hėnault *et al*. 2020; Samarasinghe *et al*. 2020; Steensels *et al*. 2021].

Bushley *et al*. [2013] detected the full sequences of the MAT1-1-1, MAT1-1-2, MAT1-1-3 and MAT1-2-1 genes in single-ascospore isolates using a genome walking/tail PCR strategy. They also observed multicellular heterokaryotic hyphae and ascospores of natural *C. sinensis* with mononucleated, binucleated, trinucleated and tetranucleated structures (*cf*. Figure 3 of [Bushley *et al*. 2013]). However, they detected the alternatively spliced transcript of the MAT1-2-1 gene of *H. sinensis* Strain 1229 with an unspliced intron I that contains 3 stop codons (*cf*. Figure 2); possible technical errors of DNA contamination were ruled out because intron II was normally spliced. This transcription phenomenon indicates translational interruption of the MAT1-2-1 transcript and the production of a largely truncated MAT1-2-1 protein encoded only by exon I of the gene without the majority of the protein, which is encoded by exons II and III. This inability to produce a full-length and functional MAT1-2-1 protein resulting in dysfunctional MAT1-2-1 mating might constitute a mechanism of coupled transcriptional-translational regulation in the control of *H. sinensis* reproduction resulting in the induction of self-sterility.

In addition to demonstrating the multicellular heterokaryotic structures of *C. sinensis* hyphae and ascospores by Bushley *et al*. [2013], their coauthors Li *et al*. [2013] obtained 15 cultures from a mono-ascospore of natural *C. sinensis*: seven homogeneous clones contained only GC-biased Genotype #1 *H. sinensis,* and eight other clones heterogeneously contained both GC-biased Genotype #1 and AT-biased Genotype #5 of *O. sinensis*. The sequences of the GC- and AT-biased genotypes of *O. sinensis* reside in independent genomes and belong to different fungi [Xiao *et al*. 2009; Zhu *et al*. 2010, 2012, 2018, 2019; Li *et al*. 2016b, 2022b; Zhu & Li 2017]. Li *et al*. [2023a, 2023b] observed two types of ascospores (fully and semiejected) of natural *C. sinensis* and found the coexistence of GC-biased Genotypes #1 and #13–14 of *O. sinensis*, AT-biased Genotypes #5–6 and #16 of *O. sinensis*, *Samsoniella hepiali* and an AB067719-type fungus. Zhang and Zhang [2015] hypothesized that the nuclei of binucleated hyphal and ascosporic cells (as well as mononucleated, trinucleated and tetranucleated cells) of natural *C. sinensis* likely contain different genetic material. Thus, these multicellular hyphal and ascosporic cells of natural *C. sinensis* might contain two or more sets of fungal genomes of independent species, which might be responsible for producing complementary mating-type proteins for sexual reproductive outcrossing.

In addition to the translated protein sequences derived from the genome assembly sequences of five *H. sinensis* strains, GenBank also lists 183 MAT1-2-1 protein sequences of various *H. sinensis* strains, which were not obtained directly through protein purification and amino acid sequencing but instead derived from the nucleotide sequences of the MAT1-2-1 gene or transcript of *H. sinensis* [Zhang *et al*. 2009, 2011, 2014; Bushley *et al*. 2013; Hu *et al*. 2013; Zhang & Zhang 2015]. These protein sequences show 98.4–100% similarity to the translated protein sequence of the transcriptome assembly GCQL01020543 of Strain L0106 [Liu *et al*. 2015]. Zhang and Zhang [2015] reported 4.7% and 5.7% allelic variations in the coding sequences of the *H. sinensis* MAT1-1-1 and MAT1-2-1 genes, respectively, and predicted that these high rates of genetic variation might cause translation disturbance of the coding sequences or 5.9% and 5.6% variations in the amino acid sequences.

As mentioned above, the MAT1-1-1 gene is not expressed in *H. sinensis* Strains CS2, L0106 and SCK05-4-3 and is even absent in the genomes of many other *H. sinensis* strains (*cf*. Tables 1 and 3) [Zhang *et al*. 2009, 2011, 2014; Liu *et al*. 2015; Zhang & Zhang 2015]. In addition to the absence of the MAT1-2-1 gene in many *H. sinensis* strains (genetic regulation of the mating process) (*cf*. Tables 1 and 3), the alternatively spliced transcript of the MAT1-2-1 gene with unspliced intron I and spliced intron II in *H. sinensis* Strain 1229 provides a translation template for the production of a largely truncated and dysfunctional MAT1-2-1 protein encoded by exon I but not by exons II and III, which constitutes an example of coupled transcriptional-translational regulation of the mating process [Bushley *et al*. 2013]. To date, there have been no experimental reports on the parallel production and direct amino acid sequencing of the mating proteins of both the *MAT1-1* and *MAT1-2* idiomorphs within a pure culture of *H. sinensis*, although many papers have used the word "expression" to describe the transcription of mating-type genes without considering other aspects of gene expression, such as epigenetic, posttranscriptional, translational and posttranslational modifications and protein activation/degradation processes.

Based on the differential occurrence and transcription of the mating-type genes, regardless of whether *H. sinensis* (Genotype #1), the postulated sole anamorph of *O. sinensis*, might be monoecious or dioecious, there may be two or more *H. sinensis* populations capable of producing either of the mating proteins of the *MAT1-1* and *MAT1-2* idiomorphs and functioning reciprocally as sexual partners for successful physiological heterothallism crossing. If this assumption is correct, the sexual partners might possess indistinguishable *H. sinensis*-like morphological and growth characteristics [Engh 1999; Kinjo & Zang 2001; Chen *et al*. 2004, 2011; Zhang *et al*. 2009; Li *et al*. 2013, 2016b; Mao *et al*. 2013]. For instance, the indistinguishable *H. sinensis* Strains 1229 and L0106 produce complementary transcripts of the mating-type genes and mating proteins of the *MAT1-1* and *MAT1-2* idiomorphs, as well as the a- and α-pheromone receptor genes, which are differentially transcribed and produce variable receptor proteins in the indistinguishable *H. sinensis* Strains Co18 and L0106. If the physiological heterothallism hypothesis is incorrect for *O. sinensis*, one of the mating proteins might be produced by heterospecific fungal species, which would result in plasmogamy and the formation of heterokaryotic cells (*cf*. Figure 3 of [Bushley *et al*. 2013]) to ensure a successful hybridization process if the heterospecific species are able to break interspecific reproduction isolation, similar to many cases of fungal hybridization that probably facilitate adaptation to the extremely adverse ecological environment on the Tibet-Qinghai Plateau [Pfennig 2007; Du *et al*. 2020; Hėnault *et al*. 2020; Samarasinghe *et al*. 2020; Steensels *et al*. 2021]. Alternatively, to complete physiological heterothallism or hybridization reproduction, mating partners might exist in three-dimensionally adjacent hyphal cells, which might make their mating choices and communicate with each other through a mating signal-based transduction system of pheromones and pheromone receptors and form "H"-shaped crossings of multicellular hyphae, as observed by Hu *et al*. [2013], Bushley *et al*. [2013] and Mao *et al*. [2013]. In particular, Mao *et al*. [2013] reported the observation of "H"-shaped morphology in *C. sinensis* hyphae that contained either AT-biased Genotype #4 or #5 of *O. sinensis* without the co-occurrence of GC-biased Genotype #1 *H. sinensis,* and the AT-biased *O. sinensis* genotypes shared indistinguishable *H. sinensis*-like morphological and growth characteristics.

Sexual Reproduction Strategy during the Lifecycle of Natural C. sinensis

The differential occurrence and transcription of the mating-type genes of both the *MAT1-1* and *MAT1-2* idiomorphs have also been observed in natural and cultivated *C. sinensis*, which contain multiple genotypes of *O. sinensis* and numerous fungal species [Engh 1999; Kinjo & Zang 2001; Chen *et al*. 2004, 2011; Xiao *et al*. 2009; Zhang *et al*. 2009, 2018; Zhu *et al*. 2010, 2012; Gao *et al*. 2011, 2012; Li *et al*. 2013, 2016b, 2016b, 2019, 2020, 2022b, 2023a, 2023b; Mao *et al*. 2013; Wei *et al*. 2016; Zhu & Li 2017; Zhong *et al*. 2018; Zhao *et al*. 2020; Yang *et al*. 2021]. As described above, transcriptome and metatranscriptome studies have shown differential transcription of mating-type genes in different maturation stages of natural and cultivated *C. sinensis* (*cf*. Figure 1), with a wide range of similarities compared with the sequences of the *H. sinensis* genes and transcripts and the unassembled metatranscriptome sequence reads, possibly indicating heterogeneous fungal sources of the transcripts. The middle and N-terminal truncated MAT1-1-1 proteins observed in natural *C. sinensis* exhibit variable hydrophobicity and alterations in the α-Helix, β-Sheet, β-Turn and Coil (*cf*. Figures 3–5, Table 4), suggesting heteromorphic folding and altered primary and secondary structures of the MAT1-1-1 proteins, which could result in dysfunctional or anomalous fungal mating processes and may indicate the heterospecific fungal sources of the proteins needed for accomplishing hybridization in natural *C. sinensis*. In addition, the variable α-pheromone receptor proteins observed in natural *C. sinensis* exhibit changes in hydrophobicity and in the α-Helix, β-Sheet, β-Turn and Coil of the proteins (*cf*. Figures 9–11, Table 6), indicating altered primary and secondary structures of the α-pheromone receptor proteins that could result in altered functionality in the sexual signal reception from mating partners in natural *C. sinensis*. Although the coexistence of larval tissues and fungal mycelia in the caterpillar body of natural *C. sinensis* indicates that larval tissues are not just culture medium that passively provide nutrients for fungal growth [Li *et al*. 2022b], the impact of host immunological reactions on the mating-type gene transcription of *O. sinensis* has not been explored, which represents the key process in the production of *O. sinensis* fruiting bodies and in the sexual life of natural and cultivated *C. sinensis*.

Regardless of whether *H. sinensis* (Genotype #1 of *O. sinensis*) is monoecious or dioecious, the sexual reproductive process of *O. sinensis* might require mating partners with the same or a different genotype of *O. sinensis* or even another fungal species to produce complementary mating proteins for physiological heterothallism or hybridization outcrossing. Thus, *O. sinensis* might have more than one monoecious or dioecious anamorph to accomplish the sexual reproduction of *O. sinensis* and the lifecycle of natural *C. sinensis*. These fungal partners might stay within a single heterokaryotic hyphal and ascosporic cell with mononucleated, binucleated, trinucleated or tetranucleated structures [*cf*. Figure 3 of Bushley *et al*. 2013], which would suggest that the following scientific observations from prior studies need to be reassessed:

(1) Li *et al*. [2013] detected GC-biased *H. sinensis* (Genotype #1) and AT-biased Genotype #5 of *O. sinensis* in eight of 15 cultures from mono-ascospores of natural *C. sinensis*. The authors misinterpreted all AT-biased genotypes as the “ITS pseudogene” components of the *H. sinensis* genome, while AT-biased Genotypes #4, #6 and #15–17 were not detected by Li *et al*. [2013] in the cultures of the mono-ascospores and the sequences of all AT-biased genotypes residing not in the genomes of GC-biased *H. sinensis* Strains 1229, CC1406-203, Co18, IOZ07 and ZJB12195 but instead in the genomes of independent *O. sinensis* fungi [Xiao *et al*. 2009; Hu *et al*. 2013; Li *et al*. 2016a, 2016b, 2022a, 2022b, 2023a, 2023b, 2024; Zhu & Li 2017; Zhu *et al*. 2018, 2019; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020].

(2) Zhu *et al*. [2010, 2012] reported the cooccurrence of GC-biased Genotypes #1 and #2 of *O. sinensis* in the stromata of natural *C. sinensis*. The ITS sequences of these two genotypes share 94.7% sequence similarity [Li *et al*. 2016b, 2022b, 2023b; Zhu & Li 2017]. The Genotype #2 sequences were located outside the phylogenetic clade of Genotype #1 in the Bayesian trees and did not reside in the genome of Genotype #1 *H. sinensis* (Li *et al*. 2022b, 2023b, 2024). The abundances of the two genotypes undergo dynamic alterations in an asynchronous, disproportional manner in the stroma of natural *C. sinensis* during maturation (Figure 12) [Zhu *et al*. 2012; Li *et al*. 2022b], indicating the genomic independence of the two GC-biased genotypes as evidence of independent *O. sinensis* fungi.

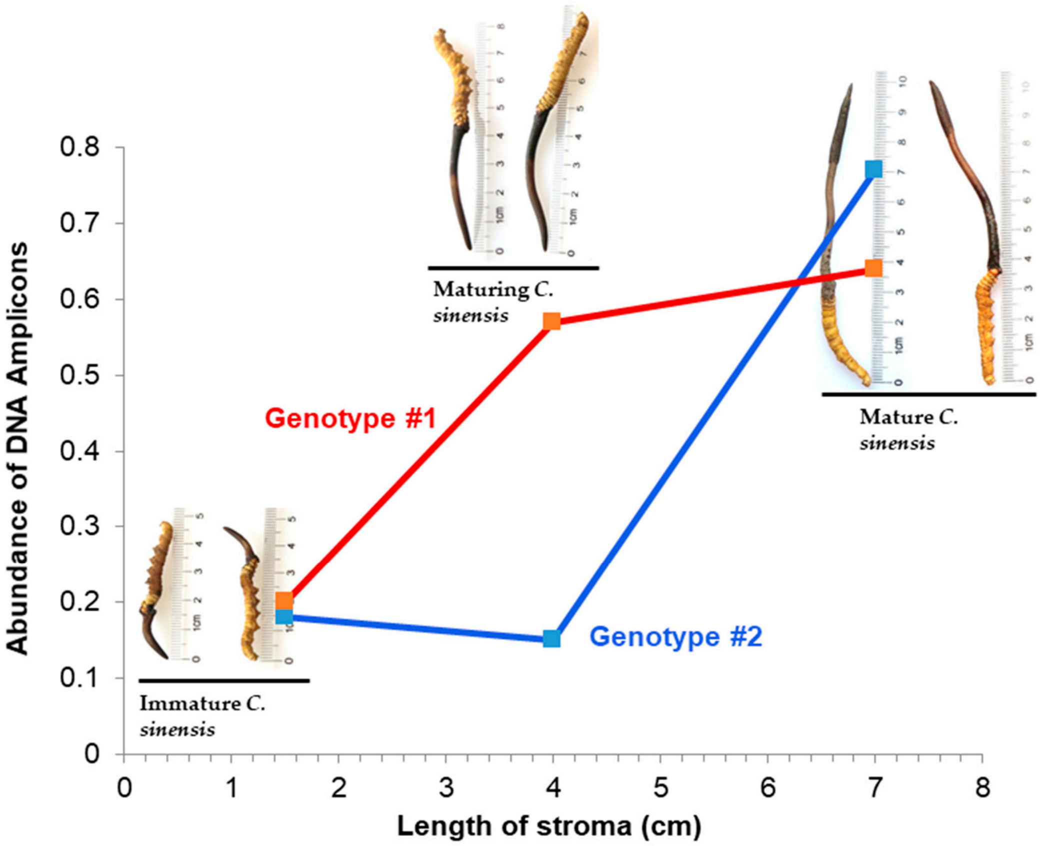


Figure 12. Dynamic alterations in the abundance of the amplicons of GC-biased Genotypes #1 and #2 of *O. sinensis* in the stromata of natural *C. sinensis* during maturation (modified from Figure 6 of [Zhu *et al*. 2012]).

(3) Chen *et al*. [2011] reported the detection of the Genotype #1 *H. sinensis* sequence AJ488255 from the caterpillar body of a natural *C. sinensis* specimen (#H1023) collected from Qinghai Province in China and the Genotype #7 sequence AJ488254 with multiple transversions and transition point mutations from the stroma of the same specimen [Li *et al*. 2016b, 2022b, 2023b; Zhu & Li 2017]. The GC-biased Genotype #7 sequence is located within the phylogenetic clade of GC-biased Genotype #1 in the Bayesian trees but does not reside in the genome of GC-biased Genotype #1 *H. sinensis* (Li *et al*. 2022b, 2023b, 2024).

(4) The cooccurrence of multiple AT-biased genotypes of *O. sinensis* and GC-biased Genotype #1 *H. sinensis* of *O. sinensis* in different combinations has been observed in the stroma, caterpillar body, ascocarps and ascospores of natural *C. sinensis* [Xiao *et al*. 2009; Zhu *et al*. 2010, 2012; Gao *et al*. 2011, 2012; Yao *et al*. 2011; Zhu & Li 2017; Li *et al*. 2022b, 2023a, 2023b]. The sequences of the mutant *O. sinensis* genotypes do not reside in the genome of GC-biased *H. sinensis* but instead belong to the genomes of independent fungi [Xiao *et al*. 2009; Zhu *et al*. 2010, 2012, 2019; Gao *et al*. 2011, 2012; Yao *et al*. 2011; Hu *et al*. 2013; Li *et al*. 2016a, 2016b, 2022b, 2023a, 2023b, 2024; Zhu & Li 2017; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020]. The abundances of the AT- and GC-biased genotypes of *O. sinensis* undergo dynamic alterations in an asynchronous, disproportional manner in the caterpillar body and stroma of *C. sinensis* during maturation, with a consistent predominance of the AT-biased genotypes, not GC-biased Genotype #1 *H. sinensis*, in the stromata [Zhu *et al*. 2010, 2012; Gao *et al*. 2011, 2012; Li *et al*. 2016b, 2022b; Zhu & Li 2017].

(5) Mao *et al*. [2013] identified AT-biased Genotype #4 or #5 of *O. sinensis* fungus without the co-occurrence of GC-biased *H. sinensis* in natural *C. sinensis* specimens collected from production areas in geographically remote locations. They also observed that the AT-biased mutant genotypes shared indistinguishable *H. sinensis*-like morphologic and growth characteristics and were able to form "H"-shaped hyphal crossings and anastomoses during germination, which are related to the sexual

reproduction of *O. sinensis*. Similarly, Kinjo and Zang [2001] reported the detection of AT-biased Genotype #4 or #5 of *O. sinensis* in several natural *C. sinensis* specimens collected from remote production areas and GC-biased Genotype #1 *H. sinensis* in other *C. sinensis* specimens collected from different production areas.

(6) Hu *et al*. [2013] reported that *H. sinensis* has hermaphroditic and homothallic characteristics. Forty larvae of Hepialidae sp. were inoculated with a mixture of two pure *H. sinensis* strains, Co18 and QH195-2. Fungal inoculation induced death and mummification of the larvae but failed to induce the development of fruiting bodies and ascospores, indicating biological separation of the larval death/mummification process and fungal fruiting body development process. The authors cited two other studies [Holliday & Cleaver 2008; Stone 2010] and commented that inoculation of ghost moth larvae of the Hepialidae family with pure *H. sinensis* consistently failed to produce fruiting bodies and ascospores. Zhang *et al*. [2013] (coauthors of [Hu *et al*. 2013]) summarized 40 years of experience in artificial cultivation of *C. sinensis* and concluded that "it is very difficult in our laboratory to induce development of the *C. sinensis* fruiting bodies, either on culture medium or on insects."

(7) Wei *et al*. [2016] reported a species contradiction between anamorphic inoculants (3 strains of GC-biased Genotype #1 *H.* sinensis: 130508-KD-2B, 20110514 and H01-20140924-03) and the sole teleomorph of AT-biased Genotype #4 of *O. sinensis* in the fruiting body of cultivated *C. sinensis*. In addition, Figure 6 of [Wei *et al*. 2016] shows two phylogenetically distinct teleomorphs of *O. sinensis*: AT-biased Genotype #4 of *O. sinensis* in cultivated *C. sinensis* and GC-biased Genotype #1 in the natural *C. sinensis* specimen G3, which was used as the teleomorphic reference in the phylogenetic analysis. Because the sequences of AT- and GC-biased genotypes of *O. sinensis* reside in independent genomes of different fungi [Xiao *et al*. 2009; Hu *et al*. 2013; Zhu *et al*. 2010, 2012, 2018, 2019; Li *et al*. 2016a, 2016b, 2020a, 2022b, 2024; Jin *et al*. 2020; Liu *et al*. 2020; Shu *et al*. 2020], Wei *et al*. [2016] demonstrated two distinct teleomorphs of *O. sinensis* and questioned the true causal fungus/fungi and anamorph-teleomorph connections of *O. sinensis* according to Koch's postulates and the sole anamorph and sole teleomorph hypotheses proposed 10 years ago by the same group of key authors [Wei *et al*. 2006].

(8) *Tolypocladium sinense* in natural *C. sinensis* was first identified and published by Li [1988]. It was subsequently isolated from natural *C. sinensis* and characterized morphologically and genetically [Jiang & Yao 2003; Chen *et al*. 2004; Leung *et al*. 2006]. Engh [1999] reported the molecular identification of the *Cordyceps*-*Tolypocladium* complex in natural *C. sinensis*. The "*Cordyceps*" sequence AJ786590 obtained by Engh [1999] was published and uploaded to GenBank by Stensrud *et al*. [2005] and phylogenetically clustered into AT-biased Group B (Genotype #4) of *O. sinensis*, along with other *C. sinensis* sequences, by Stensrud *et al*. [2007]. Barseghyan *et al*. [2011] performed a macro/micromycology study and concluded that *H. sinensis*, which is presumed to be psychrophilic, and *T. sinensis*, which is presumed to be mesophilic, are dual anamorphs of *O. sinensis*. Notably, the *O. sinensis* fungus, which has *H. sinensis*-like morphology and growth characteristics, was not genotyped molecularly in that study.

(9) Genotypes #13 (KT339190) and #14 (KT339178) of *O. sinensis* have been identified in either semiejected or fully ejected multicellular heterokaryotic ascospores collected from the same specimen of natural *C. sinensis* [Zhu & Li 2017; Li *et al*. 2022b, 2023b]. The two genotypes feature precise reciprocal substitutions of large DNA segments due to chromosomal intertwining interactions and genetic material recombination between two parental fungi, Genotype #1 *H. sinensis* (Group A by Stensrud *et al*. [2007]) and an AB067719-type Group E fungus [Zhu & Li 2017; Li *et al*. 2022b, 2023b]. A pure culture of the AB067719-type fungus has not been obtained, and its taxonomic position is unclear. More than 900 sequences highly homologous to AB067719, including *Alternaria* sp., *Ascomycota* sp., *Aspergillus* sp., *Avena* sp., *Berberis* sp., *Colletotrichum* sp., *Cordyceps* sp., *Cyanonectria* sp., *Dikarya* sp., *Fusarium* sp., *Gibberella* sp., *Hypocreales* sp., *Juglans* sp., *Lachnum* sp., *Nectria* sp., *Nectriaceae* sp., *Neonectria* sp., and *Penicillium* sp., have been uploaded to GenBank [Li *et al*. 2023b]. Chromosomal intertwining and genetic material recombination may occur after plasmogamy and karyogamy of heterospecific parental fungi under sexual reproduction hybridization or parasexuality, which is characterized by the prevalence of heterokaryosis and results in concerted

chromosome loss for transferring/substituting genetic materials without conventional meiosis [Bennett & Johnson 2003; Sherwood & Bennett 2009; Bushley *et al*. 2013; Seervai *et al*. 2013; Nakamura *et al*. 2019]. The phenomena of precise vertical transfer and reciprocal substitution of genetic materials between the chromosomes of heterospecific parental fungi that occurred differentially between the two types of ascospores collected from the same specimen of natural *C. sinensis* are distinct from the randomness and arbitrariness of horizontal environmental gene drift.

Table 7. Percent similarities between the sequences of the ITS1, 5.8S and ITS2 segments of the parental fungi (Genotype #1 *H. sinensis* AB067721 and the AB067719-type Group E fungus) and the segment sequences of the offspring variant Genotypes #13 (KT339190, in the semiejected ascospores) and #14 (KT339178, in the fully ejected ascospores) of *O. sinensis* (adapted from [Zhu & LI 2017; Li *et al*. 2022b, 2023b]).

	Percent similarity			
	ITS1-5.8S-ITS2	**ITS1**	**5.8S**	**ITS2**
	***vs.* Genotype #13 of *O. sinensis* (KT339190)**			
Genotype #1 AB067721 *O. sinensis*	86.3%	**100%**	94.8%	64.2%
Group E AB067719 fungus	88.2%	71.5%	**100%**	**99%**
	***vs.* Genotype #14 of *O. sinensis* (KT339178)**			
Genotype #1 AB067721 *O. sinensis*	87.7%	67.9%	94.9%	**100%**
Group E AB067719 fungus	89.2%	**100%**	**100%**	71.5%

(10) *Paecilomyces hepiali* was first isolated from natural *C. sinensis* by Dai *et al*. [Qiu *et al*. 1987; Dai *et al*. 1989; Jiang & Yao 2003] and renamed *Samsoniella hepiali* by Wang *et al*. [2020]. A close association of psychrophilic *H. sinensis* and mesophilic *S. hepiali* has been found in the caterpillar body, stroma, stromal fertile portion densely covered with ascocarps, and ascospores of natural *C. sinensis* and even in the formation of a fungal complex in "pure" *H. sinensis* strains that were isolated from natural *C. sinensis* and provided as gifts by a well-known mycology taxonomist [Jiang & Yao 2003; Chen *et al*. 2004; Zhu *et al*. 2007, 2010, 2012; Yang *et al*. 2008; Yao *et al*. 2011; Zhu & Wu 2015; Li *et al*. 2016b, 2016c, 2023b; Zhu & Li 2017]. Whether certain strains of these two fungal species would select each other as sexual partners will depend on their mating choices for hybridization and their abilities to break interspecific isolation barriers to adapt to extremely harsh ecological environments on the Qinghai–Tibet Plateau and the seasonal change from the extremely cold winter when *C. sinensis* is in its asexual growth phase and to the spring and early summer when *C. sinensis* switches to the sexual reproduction phase [Pfennig 2007; Du *et al*. 2020; Hėnault *et al*. 2020; Samarasinghe *et al*. 2020; Steensels & Gallone 2021].

Conclusions

This paper analyses the differential occurrence and transcription of the mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs and pheromone receptor genes of *H. sinensis*. The genomic and transcriptomic evidence is inconsistent with the self-fertilization hypothesis under homothallism and pseudohomothallism for *H. sinensis* but instead suggests self-sterility in *O. sinensis*, which utilizes physiological heterothallism or fungal hybridization strategies for sexual reproduction in natural and cultivated *C. sinensis*. *H. sinensis* Strains L0106 and 1229, either monoecious or dioecious, differentially transcribed the mating-type genes of the *MAT1-1* and *MAT1-2* idiomorphs and might become sexual partners for physiological heterothallic outcrossing. *H. sinensis* Strain L0106 possesses the transcript of an α-pheromone receptor gene but not an a-pheromone receptor gene, indicating its ability to receive a mating signal from an a-pheromone secreted by a-cells of another fungus, either the same or different fungal species, to accomplish heterothallism or hybridization outcrossing for the development and maturation of the fruiting body, ascocarps and ascospores of natural *C. sinensis*. The mutations of MAT1-1-1 and α-pheromone receptor proteins observed in natural *C. sinensis* result in dramatically altered hydrophobicity properties and secondary protein structures, probably suggesting the heterogeneity of fungal source(s) of the proteins in the natural *C. sinensis* insect-fungal complex and sexual reproduction of *O. sinensis* under heterothallism or hybridization.

Author Contributions: Conceptualization, XZL, YLL and JSZ; methodology, JSZ; formal analysis, JSZ; data curation, XZL and JSZ; writing–original draft preparation, JSZ; writing—review and editing, XZL., YLL and JSZ; project administration, YLL; funding acquisition, YLL. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant from the Science and Technology Department of Qinghai Province, China, grant number 2021-SF-A4, "Study on key technologies of conservation of natural resource and industrial upgrading of *Cordyceps sinensis*", the major science and technology projects in Qinghai Province.

Institutional Review Board Statement: Not applicable because this paper is public data reanalysis.

Informed Consent Statement: Not applicable because this paper is public data reanalysis.

Data Availability Statement: All sequence data are available in the GenBank database, except for the metatranscriptome sequences of natural *C. sinensis* that were submitted to the depository database www.plantkingdomgdb.com/Ophiocordyceps_sinensis/data/cds/Ophiocordyceps_sinensis_CDS.fas by Xia *et al*. [2017], which is currently inaccessible, but a previously downloaded cDNA file was used for the metatranscriptomic analysis.

Acknowledgments: The authors are grateful to Prof. Mu Zang, Prof. Ru-Qin Dai, Prof. Ping Zhu, Prof. Zong-Qi Liang for consultations and assistance.

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

References

1. Barseghyan GS, Holliday JC, Price TC, Madison LM, Wasser SP. Growth and cultural-morphological characteristics of vegetative mycelia of medicinal caterpillar fungus *Ophiocordyceps sinensis* G.H. Sung *et al*. (Ascomycetes) Isolates from Tibetan Plateau (P.R.China). Intl. J. Med. Mushrooms. 2011;13(6): 565–581. DOI: 10.1615/intjmedmushr.v13.i6.90
2. Bennett RJ, Johnson AD. Completion of a parasexual cycle in *Candida albicans* by induced chromosome loss in tetraploid strains. EMBO J. 2003;22(10): 2505–2515. DOI: 10.1093/emboj/cdg235
3. Bushley KE, Li Y, Wang W-J, Wang X-L, Jiao L, Spatafora JW, Yao Y-J. Isolation of the MAT1-1 mating type idiomorph and evidence for selfing in the Chinese medicinal fungus *Ophiocordyceps sinensis*. Fungal Biol. 2013;117(9): 599–610. DOI: 10.1016/j.funbio.2013.06.001
4. Chen C-S, Hseu R-S, Huang C-T. Quality control of *Cordyceps sinensis* teleomorph, anamorph, and Its products. Chapter 12, *in* (Shoyama, Y., Ed.) *Quality Control of Herbal Medicines and Related Areas*. InTech, Rijeka, Croatia (2011). pp. 223–238. www.intechopen.com
5. Chen Y-Q, Hu B, Xu F, Zhang W-M, Zhou H, Qu L-H. Genetic variation of *Cordyceps sinensis*, a fruit-body-producing entomopathegenic species from different geographical regions in China. FEMS Microbiol. Lett. 2004;230: 153–158. DOI: 10.1016/S0378-1097(03)00889-9
6. Chou PY, Fasman GD. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. Relat. Areas. Mol. Biol. 1978;47: 45–148. DOI: 10.1002/9780470122921.ch2
7. Dai R-Q, Lan J-L, Chen W-H, Li X-M, Chen Q-T, Shen C-Y. Discovery of a new fungus *Paecilomyces hepiali* Chen & Dai. Acta Agricult. Univ. Pekin. 1989;15(2): 221–224.
8. Debuchy R, Turgeo BG. Mating-Type Structure, Evolution, and Function in Euascomycetes. In (eds. Kües U. & Fischer, R.) Growth, Differentiation and Sexuality. Springer (2006). pp. 293–323.
9. Deleage G, Roux B. An algorithm for protein secondary structure prediction based on class prediction. Protein Engineering, Design and Selection 1987;1: 289–294. https://doi.org/10.1093/protein/1.4.289
10. Dong Y-Z, Zhang L-J, Wu Z-M, Gao L, Yao Y-S, Tan N-Z, Wu J-Y, Ni L-Q, Zhu J-S. Altered proteomic polymorphism in the caterpillar body and stroma of natural *Cordyceps sinensis* during maturation. PLoS ONE. 2014;9(10): e109083 DOI: 10.1371/journal.pone.0109083
11. Du X-H, Wu D-M, Kang H, Wang H-C, Xu N, Li T-T, Chen K-L. Heterothallism and potential hybridization events inferred for twenty-two yellow morel species. IMA Fungus. 2020;11: 4. https://doi.org/10.1186/s43008-020-0027-1
12. Engh IB. Molecular phylogeny of the *Cordyceps-Tolypocladium* complex. Candidate scientific thesis, Department of Biology, University of Oslo, 1999.
13. Gao L, Li X-H, Zhao J-Q, Lu J-H, Zhao J-G, Zhu J-S. Maturation of *Cordyceps sinensis* associates with alterations of fungal expressions of multiple *Ophiocordyceps sinensis* mutants with transition and

transversion point mutations in stroma of *Cordyceps sinensis*. J. Peking Univ. (Health Sci.). 2012;44(3): 454–463. https://pubmed.ncbi.nlm.nih.gov/22692321/

14. Gao L, Li X-H, Zhao J-Q, Lu J-H, Zhu J-S. Detection of multiple *Ophiocordyceps sinensis* mutants in premature stroma of *Cordyceps sinensis* by MassARRAY SNP MALDI-TOF mass spectrum genotyping. J. Peking Univ. (Health Sci.). 2011;43(2): 259–266. https://www.ncbi.nlm.nih.gov/pubmed/21503123
15. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. **Protein Identification and Analysis Tools on the Expasy** Server. Chapter 52 (*In*) John M. Walker (*ed*): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571–607.
16. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z-H, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, FriedmanN, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 2011;29(7): 644−130. https://doi.org/10.1038/nbt.1883
17. Hėnault M, Marsit S, Charron G, Landry CR. The effect of hybridization on transposable element accumulation in an undomesticated fungal species. eLife. 2020;9, e60474. DOI: https://doi.org/10.7554/eLife.60474
18. 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. DOI: 10.1615/IntJMedMushr.v10.i3.30
19. Hu X, Zhang Y-J, Xiao G-H, Zheng P, Xia Y-L, Zhang X-Y, St Leger RJ, Liu X-Z, Wang C-S. 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
20. Jiang Y, Yao Y-J. A review for the debating studies on the anamorph of *Cordyceps sinensis*. Mycosistema. 2003;22(1): 161–176.
21. Jin L-Q, Xu Z-W, Zhang B, Yi M, Weng C-Y, Lin S, Wu H, Qin X-T, Xu F, Teng Y, Yuan S-J, Liu Z-Q, Zheng Y-G. Genome sequencing and analysis of fungus *Hirsutella sinensis* isolated from *Ophiocordyceps sinensis*. AMB Expr. 2020;10: 105. https://doi.org/10.1186/s13568-020-01039-x
22. Jones SK, Bennett RJ. Fungal mating pheromones: choreographing the dating game. Fungal Genet. Biol. 2011;48(7): 668–676. doi: 10.1016/j.fgb.2011.04.001
23. Kinjo N, Zang M. Morphological and phylogenetic studies on *Cordyceps sinensis* distributed in southwestern China. Mycoscience. 2001;42: 567–574.
24. Leff SE, Rosenfeld MG, Evans RM. Complex transcriptional units: diversity in gene expression by alternative RNA processing. Ann. Rev. Biochem. 1986;55(1): 1091–117. doi:10.1146/annurev.bi.55.070186.005303
25. Leung P-H, Zhang Q-X, Wu J-Y. Mycelium cultivation, chemical composition and antitumour activity of a *Tolypocladium* sp. fungus isolated from wild *Cordyceps sinensis*. J. Appl. Microbiol. 2006;101: 275−283. DOI: 10.1111/j.1365-2672.2006.02930.x
26. Li C-L. A study of *Tolypocladium sinense* C.L. Li. sp. nov. and cyclosporin production. Acta Mycol. Sinica. 1988;7: 93−98.
27. Li X, Wang F, Liu Q, Li Q-P, Qian Z-M, Zhang X-L, Li K, Li W-J, Dong C-H. Developmental transcriptomics of Chinese cordyceps reveals gene regulatory network and expression profiles of sexual development-related genes. BMC Genomics. 2019;20: 337. https://doi.org/10.1186/s12864-019-5708-z
28. Li X-Z, Li Y-L, Wang Y-N, Zhu J-S. Genetic and phylogenetic features of the repetitive ITS copies in the *Hirsutella sinensis* genome and the genomically independent AT-biased *Ophiocordyceps sinensis* genotypes. BioRxiv (the preprint server for biology) February 03, 2024. https://www.biorxiv.org/content/10.1101/2022.12.09.519819v5
29. Li X-Z, Li Y-L, Yao Y-S, Xie W-D, Zhu J-S. Further discussion with Li *et al*. (2013, 2019) regarding the “ITS pseudogene hypothesis” for *Ophiocordyceps sinensis*. Mol. Phylogenet. Evol. 2020;146: 106728. https://doi.org/10.1016/j.ympev.2019.106728
30. Li X-Z, Li Y-L, Yao Y-S, Zhu J-S. Differential expression of *Hirsutella sinensis* genes and intraspecific genetic variation among *H. sinensis* strains. Biochem. Mol. Biol. J. 2022a;8(2): 58. DOI:10.36648/2471-8084.22.8.58
31. Li X-Z, Li Y-L, Zhu J-S. Differential expression of mating-type genes in *Hirsutella sinensis* and natural *Cordyceps sinensis*. Proc. 2021 Intl. Conf. Info. Technol. Biomed. Engineering, 2021, pp. 318−333. https://conferences.computer.org/icitbepub/pdfs/ICITBE2021-6JNtJiThlp10s9N5Sodvuy/009900a318/009900a318.pdf
32. Li Y, Hsiang T, Yang R-H, Hu X-D, Wang K, Wang W-J, Wang X-L, Jiao L, Yao Y-J. Comparison of different sequencing and assembly strategies for a repeat-rich fungal genome, *Ophiocordyceps sinensis*. J. Microbiol. Methods. 2016a;128: 1−6. DOI: 10.1016/j.mimet.2016.06.025
33. Li Y, Jiao L, Yao Y-J. Non-concerted ITS evolution in fungi, as revealed from the important medicinal fungus *Ophiocordyceps sinensis*. Mol. Phylogenet. Evol. 2013;68: 373−379. doi: 10.1016/j.ympev.2013.04.010
34. Li Y-L, Gao L, Yao Y-S, Wu Z-M, Lou Z-Q, Xie W-D, Wu J-Y, Zhu J-S. Altered GC- and AT-biased genotypes of *Ophiocordyceps sinensis* in the stromal fertile portions and ascospores of natural *Cordyceps sinensis*. PLoS ONE. 2023a;18(6): e0286865. https://doi.org/10.1371/journal.pone.0286865

35. Li Y-L, Li X-Z, Yao Y-S, Wu Z-M, Gao L, Tan N-Z, Lou Z-Q, Xie W-D, Wu J-Y, Zhu J-S. Differential cooccurrence of multiple genotypes of *Ophiocordyceps sinensis* in the stromata, stromal fertile portion (ascocarps) and ascospores of natural *Cordyceps sinensis*. PLoS ONE. 2023b;18(3): e0270776. DOI: 10.1371/journal.pone.0270776
36. Li Y-L, Li X-Z, Yao Y-S, Xie W-D, Zhu J-S. Molecular identification of *Ophiocordyceps sinensis* genotypes and the indiscriminate use of the Latin name for the multiple genotypes and the natural insect-fungi complex. Am J BioMed Sci. 2022b;14(3): 115–135. doi: 10.5099/aj220300115
37. Li Y-L, Yao Y-S, Xie W-D, Zhu J-S. Molecular heterogeneity of natural *Cordyceps sinensis* with multiple *Ophiocordyceps sinensis* fungi challenges the anamorph-teleomorph connection hypotheses and implementation of the Amsterdam Declaration of IMA. Am. J. Biomed. Sci. 2016b;8(2): 123–159. DOI:10.5099/AJ160200123
38. Li Y-L, Yao Y-S, Zhang Z-H, Xu H-F, Liu X, Ma S-L, Wu Z-M, Zhu J-S. Synergy of fungal complexes isolated from the intestines of *Hepialus lagii* larvae in increasing infection potency. J. Fungal Res. 2016c;14: 96–112.
39. Liang Z-Q, Han Y-F, Liang J-D, Dong X, Du W. Issues of concern in the studies of *Ophiocordyceps sinensis*. Microbiol. Chin. 2010;37(11): 1692–1697.
40. Liu J, Guo L-N, Li Z-W, Zhou Z, Li Z, Li Q, Xiaochen Bo X-C, Wang S-Q, Wang J-L, Ma S-C, Zheng J, Yang Y. Genomic analyses reveal evolutionary and geologic context for the plateau fungus *Ophiocordyceps sinensis*. Clin. Med. 2020;15: 107–119. https://doi.org/10.1186/s13020-020-00365-3
41. Liu Z-Q, Lin S, Baker PJ, Wu L-F, Wang X-R, Wu H, Xu F, Wang H-Y, Brathwaite ME, Zheng Y-G. Transcriptome sequencing and analysis of the entomopathogenic fungus *Hirsutella sinensis* isolated from *Ophiocordyceps sinensis*. BMC Genomics. 2015;16: 106–123. DOI: 10.1186/s12864-015-1269-y
42. Lu H-L, St Leger RJ. Chapter Seven - Insect Immunity to Entomopathogenic Fungi, Editors: Lovett B, St. Leger RJ. Advanc Genet, Academic Press, 2016. Vol. 94; pp. 251-85.
43. Mao X-M, Zhao S-M, Cao L, Yan X, Han R-C. The morphology observation of *Ophiocordyceps sinensis* from different origins. J. Environ. Entomol. 2013;35(3): 343–353.
44. Meng Q, Yu H-Y, Zhang H, Zhu W, Wang M-L, Zhang J-H, Zhou G-L, Li X, Qin Q-L, Hu S-N, Zou Z. Transcriptomic insight into the immune defenses in the ghost moth, *Hepialus xiaojinensis*, during an *Ophiocordyceps sinensis* fungal infection. Insect Biochem. Mol. Biol. 2015;64: 1–15. DOI: 10.1016/j.ibmb.2015.06.014
45. Nakamura N, Tanaka C, Takeuchi-Kaneko Y. Transmission of antibiotic-resistance markers by hyphal fusion suggests partial presence of parasexuality in the root endophytic fungus *Glutinomyces brunneus*. Mycol. Progress 2019;18: 453–462. https://doi.org/10.1007/s11557-018-1455-9
46. Peters C, Elofsson A. Why is the biological hydrophobicity scale more accurate than earlier experimental hydrophobicity scales? Proteins. 2014;82(9): 2190-2198. doi: 10.1002/prot.24582
47. Pfennig KS. Facultative Mate Choice Drives Adaptive Hybridization. Sci. 2007;318: 965–967. DOI: 10.1126/science.1146035
48. Qiu W-F, Chen Q-T, Qian C-R, Shen C-Y, Xu J-T, Wang L-Y, Zhang W-J. “Study on *Cordyceps sinensis*, a new species *Paecilomyces hepiali*” certificate of identification for the achievement in science and technology. In Archive of Institution of Chinese Materia Medica, Chinese Academy of Traditional Chinese Medicine (Unpublished; January 8, 1987) Docs #000021, #000044, and #000045.
49. Ren Y, Wan D-G, Lu X-M, Guo J-L. The study of scientific name discussion for TCM Cordyceps. LisShenzhen Med. Material Medica Res. 2013;24(9): 2211–2212.
50. Samarasinghe H, You M, Jenkinson TS, Xu J-P, James TY. Hybridization Facilitates Adaptive Evolution in Two Major Fungal Pathogens. Genes. 2020;11: 101. doi:10.3390/genes11010101
51. Seervai RNH, Jones SK, Hirakawa MP, Porman AM, Bennett RJ. Parasexuality and ploidy change in *Candida tropicalis*. Eukaryotic Cell. 2013;12(12): 1629–1640. DOI: 10.1128/EC.00128-13
52. Sherwood RK, Bennett RJ. Fungal meiosis and parasexual reproduction--lessons from pathogenic yeast. Current Opinion Microbiol. 2009;12(6): 599-607. DOI: 10.1016/j.mib.2009.09.005
53. Shu R-H, Zhang J-H, Meng Q, Zhang H, Zhou G-L, Li M-M, Wu P-P, Zhao Y-N, Chen C, Qin Q. A new high-quality draft genome assembly of the Chinese cordyceps *Ophiocordyceps sinensis*. Genome Biol. Evol. 2020;12(7): 1074–1079. DOI: 10.1093/gbe/evaa112
54. Simm S, Einloft J, Mirus O, Schleiff E. 50 years of amino acid hydrophobicity scales: revisiting the capacity for peptide classification. Biol. Res. 2016;49(1): 31. doi: 10.1186/s40659-016-0092-5
55. Steensels J, Gallone B, Verstrepen KJ. Interspecific hybridization as a driver of fungal evolution and Adaptation. Nat. Rev. Microbiol. 2021;19: 485–500.
56. Stensrud Ø, Hywel-Jones NL, Schumacher T. Towards a phylogenetic classification of *Cordyceps*: ITS nrDNA sequence data confirm divergent lineages and paraphyly. Mycol. Res. 2005;109: 41–56.
57. Stensrud Ø, Schumacher T, Shalchian-Tabrizi K, Svegardenib IB, Kauserud H. Accelerated nrDNA evolution and profound AT bias in the medicinal fungus *Cordyceps sinensis*. Mycol. Res. 2007;111: 409–415.
58. Stone R. Improbable partners aim to bring biotechnology to a Himalayan kingdom. Sci. 2010;327: 940–941.

59. Sung G-H, Hywel-Jones NL, Sung J-M, Luangsa-ard JJ, Shrestha B, Spatafora JW. Phylogenetic classification of *Cordyceps* and the clavicipitaceous fungi. Stud. Mycol. 2007;57: 5–59.
60. Turgeon BG, Yoder OC. Proposed nomenclature for mating type genes of filamentous ascomycetes. Fungal Genet. Biol. 2000;31: 1–5.
61. Wang J-B, St. Leger RJ, Wang C-S. Advances in Genomics of Entomopathogenic Fungi. Advan. Genetics 2016;94: 1–39. DOI: 10.1016/bs.adgen.2016.01.002
62. Wang Y, Stata M, Wang W, Stajich JE, White MM, Moncalvo JM. Comparative genomics reveals the core gene toolbox for the fungus-insect symbiosis. mBio. 2018;9: e00636-18. DOI: 10.1128/mBio.00636-18
63. Wang Y-B, Wang Y, Fan Q, Duan D-E, Zhang G-D, Dai R-Q, Dai Y-D, Zeng W-B, Chen Z-H, Li D-D, Tang D-X, Xu Z-H, Sun T, Nguyen T-T, Tran N-L, Dao V-M, Zhang C-M, Huang L-D, Liu Y-J, Zhang X-M, Yang D-R, Sanjuan T, Liu X-Z, Yang Z-L, Yu H. Multigene phylogeny of the family *Cordycipitaceae* (Hypocreales): new taxa and the new systematic position of the Chinese cordycipitoid fungus *Paecilomyces hepiali*. Fungal Diversity. 2020; 103:1–46.
64. Wei J-C, Wei X-L, Zheng W-F, Guo W, Liu R-D. Species identification and component detection of *Ophiocordyceps sinensis* cultivated by modern industry. Mycosystema. 2016;35(4): 404–410.
65. Wei X-L, Yin X-C, Guo Y-L, Shen N-Y, Wei, J.-C. Analyses of molecular systematics on *Cordyceps sinensis* and its related taxa. Mycosystema. 2006;25(2): 192–202.
66. Wilson AM, Wilken PM, van der Nest MA, Steenkamp ET, Wingfield MJ, Wingfield BD. Homothallism: an umbrella term for describing diverse sexual behaviours. IMA Fungus. 2015;6(1): 207–214.
67. Xia E-H, Yang D-R, Jiang J-J, Zhang Q-J, Liu Y, Liu Y-L, Zhang Y, Zhang H-B, Shi C, Tong Y, Kim C-H, Chen H, Peng Y-Q, Yu Y, Zhang W, Eichler EE, Gao L-Z. The caterpillar fungus, *Ophiocordyceps sinensis*, genome provides insights into highland adaptation of fungal pathogenicity. Sci. Rep. 2017;7: 1806. DOI:10.1038/s41598-017-01869-z
68. Xia F, Liu Y, Shen G-L, Guo L-X, Zhou X-W. Investigation and analysis of microbiological communities in natural *Ophiocordyceps sinensis*. Can. J. Microbiol. 2015;61: 104–111. DOI: 10.1139/cjm-2014-0610
69. Xiang L, Li Y, Zhu Y, Luo H, Li C, Xu X, Sun C, Song J-Y, Shi L-H, He L, Sun W, Chen S-L. Transcriptome analysis of the *Ophiocordyceps sinensis* fruiting body reveals putative genes involved in fruiting body development and cordycepin biosynthesis. Genomics. 2014;103: 154−159. https://doi.org/10.1016/j.ygeno.2014.01.002
70. Xiao W, Yang J-P, Zhu P, Cheng K-D, He H-X, Zhu H-X, Wang Q. Non-support of species complex hypothesis of *Cordyceps sinensis* by targeted rDNA-ITS sequence analysis. Mycosystema. 2009;28(6): 724–730.
71. Yang J-L, Xiao W, He H-H, Zhu H-X, Wang S-F, Cheng K-D, Zhu P. Molecular phylogenetic analysis of *Paecilomyces hepiali* and *Cordyceps sinensis*. Acta Pharmaceut. Sinica. 2008;43(4): 421–426.
72. Yang J-Y, Tong X-X, He C-Y, Bai J, Wang F, Guo J-L. Comparison of endogenetic microbial community diversity between wild *Cordyceps sinensis*, artificial *C. sinensis* and habitat soil. Chin. J. Chin. Materia Medica. 2021;46(12): 3106–3115.
73. Yao, Y.-S., Gao, L., Li, Y.-L., Ma, S.-L., Wu, Z.-M., Tan, N.-Z., Wu, J.-Y., Ni, L.-Q., Zhu, J.-S. Amplicon density-weighted algorithms for analyzing dissimilarity and dynamic alterations of RAPD polymorphisms of *Cordyceps sinensis*. Beijing Da Xue Xue Bao Yi Xue Bao. 2014;46: 618−628. (http://xuebao.bjmu.edu.cn/EN/Y2014/V46/I4/618)
74. Yao Y-S, Zhou Y-J, Gao L, Lu J-H, Zhu J-S. Dynamic alterations of the differential fungal expressions of *Ophiocordyceps sinensis* and its mutant genotypes in stroma and caterpillar during maturation of natural *Cordyceps sinensis*. J. Fungal Res. 2011;9(1): 37–49,53.
75. Zhang, S.; Zhang, Y.-J. Molecular evolution of three protein-coding genes in the Chinese caterpillar fungus *Ophiocordyceps sinensis*. Microbiol. China. 2015;42(8): 1549−1560.
76. Zhang S, Zhang Y-J, Liu X-Z, Wen H-A, Wang M, Liu D-S. Cloning and analysis of the MAT1-2-1 gene from the traditional Chinese medicinal fungus *Ophiocordyceps sinensis*. Fungal Biol. 2011;115: 708−714.
77. Zhang S, Zhang Y-J, Shrestha B, Xu J-P, Wang C-S, Liu X-Z. *Ophiocordyceps sinensis* and *Cordyceps militaris*: research advances, issues and perspectives. Mycosystema. 2013;32:577−597.
78. Zhang S-W, Cen K, Liu Y, Zhou X-W, Wang C-S. Metatranscriptomics analysis of the fruiting caterpillar fungus collected from the Qinghai-Tibetan plateau. Sci. Sinica Vitae. 2018;48(5): 562−570.
79. Zhang Y-J, Li E-W, Wang C-S, Li Y-L, Liu X-Z. *Ophiocordyceps sinensis*, the flagship fungus of China: terminology, life strategy and ecology. Mycol. 2012;3(1): 2–10. https://www.tandfonline.com/doi/full/10.1080/21501203.2011.654354
80. Zhang Y-J, Sun B-D, Zhang S, Wàngmŭ, Liu X-Z, Gong W-F. Mycobiotal investigation of natural *Ophiocordyceps sinensis* based on culture-dependent investigation. Mycosistema. 2010;29(4): 518–527.
81. Zhang Y-J, Xu L-L, Zhang S, Liu X-Z, An Z-Q, Wàngmŭ, *et al*. Genetic diversity of *Ophiocordyceps sinensis*, a medicinal fungus endemic to the Tibetan Plateau: implications for its evolution and conservation. BMC Evol. Biol. 2009;9: 290. doi:10.1186/1471-2148-9-290

82. Zhang Y-J, Zhang S, Li Y-L, Ma S-L, Wang C-S, Xiang M-C, Liu X, An Z-Q, Xu J-P, Liu X-Z. Phylogeography and evolution of a fungal–insect association on the Tibetan Plateau. Mol. Ecol. 2014:23: 5337−5355. DOI: 10.1111/mec.12940
83. Zhao Y-N, Zhang J-H, Meng Q, Zhang H, Zhou G-L, Li M-M, Wu P-P, Shu R-H, Gao X-X, Guo L, Tong Y, Cheng L-Q, Guo L, Chen C, Qin Q. Transcriptomic analysis of the orchestrated molecular mechanisms underlying fruiting body initiation in Chinese cordyceps. Gene. 2020;763: 145061. https://doi.org/10.1016/j.gene.2020.145061
84. Zheng P, Wang C-S. Sexuality Control and Sex Evolution in Fungi. Sci Sinica Vitae. 2013;43:1090–1097.
85. Zhong X, Gu L, Wang H-Z, Lian D-H, Zheng Y-M, Zhou S, Zhou W, Gu J, Zhang G, Liu X. Profile of *Ophiocordyceps sinensis* transcriptome and differentially expressed genes in three different mycelia, sclerotium and fruiting body developmental stages. Fungal Biol. 2018;122: 943–951. DOI: 10.1016/j.funbio.2018.05.011
86. Zhu J-S, Gao L, Li X-H, Yao Y-S, Zhou Y-J, Zhao J-Q, Zhou Y-J. Maturational alterations of oppositely orientated rDNA and differential proliferations of CG:AT-biased genotypes of *Cordyceps sinensis* fungi and *Paecilomyces hepiali* in natural *C. sinensis*. Am. J. Biomed. Sci. 2010;2(3): 217–238. doi: 10.5099/aj100300217
87. Zhu J-S, Guo Y-L, Yao Y-S, Zhou Y-J, Lu J-H, Qi Y, Chen W, Zheng T-Y, Zhang L, Wu Z-M, Zhang L-J, Liu X-J, Yin W-T. Maturation of *Cordyceps sinensis* associates with co-existence of *Hirsutella sinensis* and *Paecilomyces hepiali* DNA and dynamic changes in fungal competitive proliferation predominance and chemical profiles. J. Fungal Res. 2007;5(4): 214–224.
88. Zhu J-S, Halpern GM, Jones K. The scientific rediscovery of a precious ancient Chinese herbal regimen: *Cordyceps sinensis*: Part I. J. Altern. Complem. Med. 1998a;4(3): 289–303.
89. Zhu J-S, Halpern GM, Jones K. The scientific rediscovery of an ancient Chinese herbal medicine: *Cordyceps sinensis*: Part II. J. Altern. Complem. Med. 1998b;4(4): 429–457.
90. Zhu J-S, Li C-L, Tan N-Z, Berger JL, Prolla TA. Combined use of whole-gene expression profiling technology and mouse lifespan test in anti-aging herbal product study. Proc. 2011 New TCM Products Innovation and Industrial Development Summit, Hangzhou, China (Nov 27, 2011). pp. 443–448. https://xueshu.baidu.com/usercenter/paper/show?paperid=08341c17fa58c8f85584b92572b90f75&site=xueshu_se
91. Zhu J-S, Li Y-L. A Precious Transitional Chinese Medicine, *Cordyceps sinensis*: Multiple heterogeneous *Ophiocordyceps sinensis* in the insect-fungi complex. Lambert Academic Publishing, Saarbrüchen. Germany (2017).
92. Zhu J-S, Li Y-L, Yao Y-S, Xie W-D. Molecular evidence not supporting the use of *Hirsutella sinensis* strain EFCC7287 as the taxonomic standard for multiple *Ophiocordyceps sinensis* fungi. Intl. J. Mol. Biol. 2018;3: 114−115.
93. Zhu J-S, Li Y-L, Yao Y-S, Xie W-D. The multiple genotypes of *Ophiocordyceps sinensis* and the “ITS pseudogene” hypothesis. Mol. Phylogenet. Evol. 2019;139: 106322. doi: 10.1016/j.ympev.2018.10.034
94. Zhu J-S, Wu J-Y. Genetic heterogeneity of natural *Cordyceps sinensis* with co-existence of multiple fungi. Chin. J. Cell Biol. 2015;37(2): 284–298.
95. Zhu J-S, Zhao J-G, Gao L, Li X-H, Zhao J-Q, Lu J-H. Dynamically altered expressions of at least 6 *Ophiocordyceps sinensis* mutants in the stroma of *Cordyceps sinensis*. J. Fungal Res. 2012;10(2): 100–112.