Multiple GC-biased repetitive ITS copies in the *Hirsutella sinensis* genome are not generated by RIP mutagenesis involving transition point mutations

Xiu-Zhang Li[a], Yu-Ling Li[a], Ya-Nan Wang[b], Jia-Shi Zhu[a,c,*]

[a] State Key Laboratory of Plateau Ecology and Agriculture, Qinghai Academy of Animal and Veterinary Sciences, Qinghai University, Xining, Qinghai 810016, China

[b] State Key Laboratory of Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

[c] Institute of Biopharmaceutical and Health Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China

[*] Corresponding author:

E-mail address: zhujosh@163.com (J-S Zhu)

Short title:

H. sinensis genomic ITS repeats are not generated by RIP mutagenesis

Abstract

Li et al. (2013, 2020b) hypothesized that AT-biased genotypes of *Ophiocordyceps sinensis* are generated through "repeat-induced point mutation (RIP)" and exist as "ITS pseudogenes in a single genome" of GC-biased Genotype #1 *Hirsutella sinensis*. This study revealed multiple GC-biased repetitive ITS copies in the *H. sinensis* genome, including multiple insertion/deletion and transversion alleles, which might not be generated through RIP mutagenesis that theoretically causes C-to-T and G-to-A transitions. The repetitive ITS copies are genetically and phylogenetically distinct from the AT-biased genotypes of *O. sinensis*, which possess multiple scattered transition alleles. The sequences of GC- and AT-biased Genotypes #2–17 were absent in the genome assemblies ANOV00000000, JAAVMX000000000, LKHE00000000, LWBQ00000000 and NGJJ00000000 of *H. sinensis* strains but instead in the genomes of independent *O. sinensis* fungi. The GC- and AT-biased genotypes of *O. sinensis* were found to differentially occur in different combinations in different compartments of natural *Cordyceps sinensis*. Their abundances were dynamically altered in an asynchronous, disproportional manner during *C. sinensis* maturation. Metatranscriptomic analysis of natural *C. sinensis* revealed transcriptional silencing of the 5.8S genes of all *C. sinensis*-colonizing fungi, including *H. sinensis*. The problematic design of the 5.8S gene transcription assay and data analysis reported by Li et al. (2013) have provided controversial and incomplete evidence to support the notion that the 5.8S genes of AT-biased genotypes of *O. sinensis* are permanently nonfunctional "ITS pseudogenes". These results demonstrate the genomic independence of 17 genotypes belonging to independent *O. sinensis* fungi. In parallel with GC-biased genotypes, AT-biased *O. sinensis* genotypes might have been generated during the long history of evolution from a common ancestor through RIP or other mutagenesis mechanisms in response to the extremely harsh environment of the Qinghai-Tibet Plateau.

KEY WORDS: Multiple genotypes of *Ophiocordyceps sinensis*; genomic independence; repeat-induced point mutations (RIP); insertion/deletion, transversion and transition point mutations; 5.8S gene transcript; transcriptional silencing.

I. Introduction

Natural *Cordyceps sinensis* is one of most valued therapeutic agents in traditional Chinese medicine and has a rich history of clinical use for health maintenance, disease amelioration, postdisease/surgery recovery and antiaging therapy (Zhu et al. 1998a, 1998b; Li et al. 2001; Tan et al. 2001). The Chinese Pharmacopeia defines natural *C. sinensis* as an insect-fungi complex, which includes the remains of Hepialidae moth larva and >90 fungal species from >37 genera, with differential co-occurrence of some of 17 genotypes of *Ophiocordyceps sinensis*, *i.e.*, natural *C. sinensis* ≠ *O. sinensis* fungus (Engh 1999; Kinjo and Zang 2001; Chen et al. 2004, 2011; Stensrud et al. 2005, 2007; Wei et al. 2006, 2016; Yang et al. 2008; Xiao et al. 2009; Zhang et al. 2009, 2010; Zhu et al. 2010, 2012, 2018, 2019; Gao et al. 2011, 2012; Yao et al. 2011; Mao et al. 2013; Xia et al. 2015; Zhu and Wu 2015; Li et al. 2016b, 2020a, 2022a, 2022c, 2022d, 2022e; Zhu and Li 2017). However, the Latin name *C. sinensis* has been indiscriminately used for both the wild insect-fungi complex and the teleomorph/holomorph of the *C. sinensis* fungus since the 1840s (Yao & Zhu 2016; Li et al. 2022a). The fungus was renamed *Ophiocordyceps sinensis* using *Hirsutella sinensis* Strain EFCC 7287 (GC-biased Genotype #1) as the nomenclature reference (Sung et al. 2007), while the name of the wild and cultivated insect-fungi complex has remained unchanged. Since then, especially after the improper implementation of the IMA's One Fungus=One name nomenclature rule in *O. sinensis* research (Hawksworth et al. 2007; Zhang et al. 2013; Li et al. 2022a), the name *O. sinensis* has been used to refer to multiple teleomorphic and anamorphic fungi that have distinct genomes and has been indiscriminately applied to the natural insect-fungi complex (Zhu and Wu 2015; Yao & Zhu 2016; Li et al. 2016b, 2022a, 2022e; Zhu and Li 2017). The confusion surrounding Latin name use continues and has even spread from the scientific community to the public media and mass markets, causing a significant decrease in the wholesale and retail prices of natural *C. sinensis* and related fermentation products. Because a consensus Latin name for the wild insect-fungi product has not been identified by taxonomists and multiple genotypes of *O. sinensis* with undetermined taxonomic positions are referred to by the same Latin name (Zhang et al. 2012; Ren et al. 2013; Yao & Zhu 2016; Li et al. 2022a, 2022d, 2022e), we temporarily refer to the multiple genotype fungi as *O. sinensis*, including Genotype #1 *H. sinensis*, and we continue to refer to the wild and cultivated insect-fungi complexes as *C. sinensis*; however, this practice only partially ameliorates the academic confusion arising from the indiscriminate use of Latin names and will likely be replaced by the exclusive use of distinct Latin names in the future.

79 All 17 genotypes of *O. sinensis* are indiscriminately referred to by the same Latin name (Zhu et
80 al. 2010, 2019; Yao & Zhu 2016; Li et al. 2020a, 2022a, 2022c, 2022d, 2022e). Kinjo and Zang (2001)
81 were the first to report 2 groups and 3 clades of mutant *O. sinensis* sequences, which were mistakenly
82 believed to represent intraspecific variants based on the observation of the same *H. sinensis*-like
83 morphological and growth characteristics. Mao et al. (2013) confirmed that the fungi in Groups B–C
84 (Genotypes #4–5) shared the same *H. sinensis*-like morphological and growth characteristics.
85 Stensrud et al. (2007) used a Bayesian phylogenetic algorithm and confirmed the existence of 3 clades
86 of mutant sequences: GC-biased Group A and AT-biased Groups B–C of *O. sinensis*, in addition to 2
87 other clades (Groups D–E) believed to be unrelated to *O. sinensis*. All the mutants, including the
88 sequences in Groups D–E, were annotated as *C. sinensis* or *O. sinensis* under GenBank taxid 72228.
89 Stensrud et al. (2007) reported that the variation in the 5.8S gene sequences of *O. sinensis* in Groups
90 A–C "far exceeds what is normally observed in fungi … even at higher taxonomic levels (genera and
91 family)" that the mutant genotypes with "profound AT bias … share a common ancestor" and
92 originate from "accelerated nrDNA evolution" (nrDNA: nuclear ribosomal DNA). In addition to the
93 use of "universal" primers, Xiao et al. (2009) designed pairs of genotype-specific primers and used a
94 nested PCR technique to identify the differential coexistence of multiple *O. sinensis* mutants in the
95 stroma and caterpillar body of natural *C. sinensis* in different maturational stages; these authors
96 believed that the genetically variable sequences likely belonged to independent *O. sinensis* fungi.
97 Figure 3 of (Bushley et al. 2013) demonstrates that *C. sinensis* hyphae and ascospores are
98 heterokaryons containing multiple mono-, bi-, and trinucleated cells. Li et al. (2013) reported the
99 identification of heterogeneous ITS sequences of both GC-biased Genotype #1 and AT-biased
100 Genotype #5 of *O. sinensis* in 8 out of 15 clones following 25 days of liquid incubation of *C. sinensis*
101 mono-ascospores. They reported the identification of the 5.8S gene cDNA of Genotype #1 but not
102 Genotype #5 in a cDNA library of Strain 1220, one of the 8 heterogeneous clones. By ignoring the
103 simultaneous identification of homogenous Genotype #1 in 7 other clones derived from the same
104 mono-ascospores, these authors concluded that "Different types of ITS sequences were thus proved
105 to coexist in a single genome" and proposed the hypothesis of "ITS pseudogene … in a single
106 genome", which was extended to all AT-biased ITS genotypes. Considering the uncertain definition
107 of "ITS pseudogene … in a single genome" based on controversial evidence, Li et al. (2017) extended
108 this hypothesis to "rDNA pseudogenes", and Li et al. (2020b) further hypothesized that all AT-biased
109 *O. sinensis* genotypes "could have emerged either before or after" a new GC-biased "ITS haplotype

110 was generated" through "repeat-induced point mutations (RIP)" and co-occur as repetitive ITS copies
111 in "a single genome" of GC-biased Genotype #1 *H. sinensis*.
112 This paper reanalyzes the relevant genomic and transcriptomic sequences in the GenBank
113 database. We discuss the genetic and phylogenetic characteristics of multiple repetitive ITS copies in
114 the genome of Genotype #1 of *O. sinensis*, the genomic residency of AT-biased genotypes in the
115 genome of Genotype #1 *H. sinensis*, and whether sufficient evidence exists to reach a conclusion that
116 the 5.8S genes of AT-biased genotypes are permanently nonfunctional pseudogenes and those of GC-
117 biased genotypes are functional counterparts.
118

119 II. Materials and Methods

120 II-1. Reagents

121 Common laboratory reagents such as agarose, electrophoresis reagents, nylon membranes, and
122 blocking solution, etc., were purchased from Beijing Bioland Technology Company. Endonuclease
123 *Ava*I, *Dra*I, and *Eco*RI are products of New England BioLabs, United States. Mercuric chloride (0.1%)
124 for surface sterilization of freshly collected *C. sinensis* specimens was a gift from the Institute of
125 Microbiology, Chinese Academy of Sciences. The Universal DNA Purification kit was a product of
126 TIANGEN BIOTECH Company, China. The DNeasy Plant Mini Kit was a product of Qiagen
127 Company, Germany. The Gel Extraction Kit was a product of Omega Bio-Tek, United States. The
128 *Taq* PCR reagent kit and Vector NTI Advance 9 software were purchased from Invitrogen, United
129 States. All primers were synthesized by Invitrogen Beijing Lab. or Beijing Bomaide Technology Co.
130 The DIG High Prime DNA Labeling and Detection Kit was purchased from Roche Diagnostics,
131 Indiana, USA.
132

133 II-2. Information on fungal species in natural *C. sinensis*

134 Information regarding intrinsic fungal species in natural *C. sinensis* was summarized by Jiang
135 and Yao (2013), and subsequent discoveries of additional species among natural *C. sinensis*
136 specimens were reviewed by Li et al. (2016b, 2022a, 2022e). The sequences of 17 genotypes of *O.*
137 *sinensis* have been obtained and uploaded to GenBank by many molecular mycology groups since
138 1999. A BLAST search of the GenBank database revealed 668 ITS sequences of *O. sinensis*, which

139 were uploaded. These genotypic sequences were analyzed genetically and phylogenetically by
140 Stensrud et al. (2007), Zhu et al. (2010, 2012), Zhang et al. (2013), Zhu & Wu (2015), Li et al. (2016b,
141 2022a, 2022d. 2022e), and Zhu & Li (2017).

142

143 II-3. Genetic, genomic and transcriptomic sequences from GenBank

144 Five genome assemblies, ANOV00000000, JAAVMX000000000, LKHE00000000,
145 LWBQ00000000 and NGJJ00000000, for *H. sinensis* Strains Co18, IOZ07, 1229, ZJB12195 and
146 CC1406-203, respectively, were uploaded to GenBank (Hu et al. 2013; Shu et al. 2020; Li et al. 2016a;
147 Jin et al. 2020; Liu et al. 2020) and used for the analysis of repetitive ITS copies and multiple GC-
148 and AT-biased genotypes.

149 mRNA transcriptome assembly GCQL00000000, for *H. sinensis* Strain L0106, was uploaded to
150 GenBank by Liu et al. (2015). Two metatranscriptome studies of natural *C. sinensis* have been
151 reported. Xiang et al. (2014) uploaded the GAGW00000000 metatranscriptome assembly to
152 GenBank, which was obtained from natural specimens (unknown maturational stages) collected from
153 Kangding County of Sichuan Province, China. Xia et al. (2017) uploaded the unassembled
154 SRR5428527 Shotgun Read Archive (SRA) to GenBank for mature *C. sinensis* specimens collected
155 from Deqin County of Yunnan Province, China, but the assembled metatranscriptome sequences were
156 uploaded to www.plantkingdomgdb.com/Ophiocordyceps_sinensis/data/cds/Ophiocordyceps_sinensis_CDS.fas,
157 which is currently inaccessible, but the file has been downloaded for analysis.

158

159 II-4. Genetic and phylogenetic analyses

160 The ITS sequences of multiple genotypes of *O. sinensis* and multiple genomic repetitive ITS
161 copies as well as metatranscriptomic/transcriptomic sequences were analyzed using the GenBank
162 BLAST algorithms (https://blast.ncbi.nlm.nih.gov/blast/). A Bayesian phylogenetic tree of the ITS
163 sequences was inferred using MrBayes v3.2.7 software (Markov chain Monte Carlo [MCMC]
164 algorithm) based on $3x10^4$ samples, with a sampling frequency of 10^3 iterations after discarding the
165 first 25% of samples from a total of $4x10^6$ iterations (Huelsenbeck and Ronquist 2001; Li et al. 2022a,
166 2022d); this phylogenetic analysis was performed at Nanjing Genepioneer Biotechnologies Co.

167

168 ## II-5. Collection and on-site processing of fresh *C. sinensis*

169 Fresh *C. sinensis* specimens were purchased in a local market in Kangding County of Sichuan
170 Province, China (3800–4600 m' altitude; latitude 30°049'N, longitude 101°959'E) (Zhu et al. 2010;
171 Yao et al. 2014). Immature *C. sinensis* specimens were characterized by a plump caterpillar body and
172 very short stroma (1.0–2.0 cm). Maturing *C. sinensis* specimens were characterized by a plump
173 caterpillar body and stroma of 2.5–5.0 cm in length, without the formation of an expanded fertile
174 portion close to the stromal tip. Mature *C. sinensis* specimens were characterized by a plump
175 caterpillar body and long stroma (>5.0 cm) and by the formation of an expanded fertile portion close
176 to the stromal tip, which was densely covered with numerous ascocarps (Zhu et al. 2010; Li et al.
177 2022c). Governmental permission was not required for *C. sinensis* purchases in local markets, and
178 the collections of *C. sinensis* specimens from sales by local farmers fall under the governmental
179 regulations for traditional Chinese herbal products.

180 Freshly collected *C. sinensis* specimens were washed thoroughly in running water with gentle
181 brushing and surface sterilized in 0.1% mercuric chloride for 10 min, followed by washing 3 times
182 with sterile water (Zhu et al. 2010; Li et al. 2022c). The thoroughly cleaned and sterilized *C. sinensis*
183 specimens were immediately frozen in liquid nitrogen and kept frozen during transportation and
184 storage prior to further processing.

185

186 ## II-6. Extraction of genomic DNA and preparation of *C. sinensis* DNA

187 Caterpillar bodies and stromata of sterilized, frozen *C. sinensis* were separately ground into
188 powder in liquid nitrogen. Genomic DNA was individually extracted from the powder samples using
189 a DNeasy Plant Mini Kit.

190

191 ## II-7. PCR amplification and amplicon sequencing

192 Genomic DNA of immature, maturing and mature stromata of natural *C. sinensis* was used as the
193 template for PCR amplification. *H. sinensis*-specific primers were designed (a forward primer *Hsprp1*:
194 5' ATTATCGAGTCACCACTCCCAAACCCCC 3'; and a reverse primer *Hsprp3*: 5'

195 AGGTTCTCAGCGAGCTA 3') based on alignment analysis of the ITS sequences (AB067721,
196 AB067740, and AB067744 representing Genotypes #1 and #4–5 of *O. sinensis*, respectively).

197 A touch-down PCR protocol was used, in which the annealing temperature was initially set at
198 70 °C and was then decreased stepwise with a decrement of 0.3 °C every cycle, for a total of 36 cycles
199 (Zhu et al. 2007, 2010). PCR amplicons were subjected to agarose gel electrophoresis, and DNA
200 bands were recovered, purified using a gel extraction kit, and sequenced.

201

202 II-8. Restriction digestion assay and genotypic density analysis

203 The target PCR moieties (440 bp) based on the design of the primers (*Hsprp1* and *Hsprp3*) from
204 sequences of GC- and AT-genotypes of *O. sinensis* were subjected to overnight digestion with
205 endonuclease *Eco*RI, which is specific to the GC-biased Genotype #1 sequence (*Eco*RI-sensitive; *cf.*
206 Figure 1 below). AT-biased genotypes are *Eco*RI-resistant due to a single-base mutation that results
207 in the loss of the *Eco*RI enzymatic site. The *Eco*RI digestion products were subjected to agarose gel
208 electrophoresis, and the digestion fragments were recovered from the gel and sequenced. DNA
209 sequences were analyzed with Vector NTI Advance 9.

210 The density of PCR amplicons and the *Eco*RI digestion fragments in agarose gels were analyzed
211 and quantified using a gel imaging system with Quantity One software (Universal Hood II Molecular
212 Imager Gel Doc XR System, Bio-Rad, CA, USA). To monitor the alterations in *O. sinensis* genotypes
213 during *C. sinensis* maturation, the density of each DNA band in different agarose gels was normalized
214 to the density of the target 440-bp moiety that served as an internal control.

215

216 II-9. Preparation of fungus-specific ITS1 probe and 18S internal control probe

217 An *H. sinensis*-specific probe (158 bp) was designed based on the ITS1 sequence of GC-biased
218 Genotype #1 *H. sinensis* and generated through PCR using a forward *Hsprp1* primer and a reverse
219 *Hsprp2* primer (5' ATTTGCTTGCTTCTTGACTGAGAGATGCC 3').

220 A nonspecific 18S internal control probe was designed based on alignments of GenBank 18S
221 rDNA sequences (DQ838795, AB187268, AB067735, AB067736 and AB080090) of several *C.*
222 *sinensis*-associated fungi. A highly homologous region of 18S rDNA, located over 700 bp upstream

223 of the ITS1 regions of *C. sinensis*-associated fungi, was selected. The 18S control probe was
224 generated through PCR using a forward *Icprp1* primer (5' ACCGTCGTAGTCTTAACCAT 3') and
225 a reverse *Icprp2* primer (5' ATCGGCTTGAGCCGATAGTC 3').

226 Both amplicons were confirmed by sequencing, and the probes were labeled using the DIG High
227 Prime DNA Labeling and Detection Kit.

228

229 **II-10. Southern blotting analysis**

230 The stromata and caterpillar bodies of immature and mature *C. sinensis* were used for Southern
231 blotting. Genomic DNA was extracted from the samples and prepared by endonuclease digestion
232 using *Ava*I, *Dra*I, and *Eco*RI. The digested genomic DNA (5 μg of each sample) was subjected to
233 agarose gel electrophoresis and transferred onto a nylon membrane through capillary transfer,
234 followed by baking the membrane at 58 °C for 4 hours for fixation, blocking the binding sites on the
235 membrane in blocking solution, and hybridizing with the labeled *H. sinensis*-specific probe at 40 °C
236 overnight. The hybridized transblotted membrane was viewed by colorimetric photography using a
237 gel imaging system (Bio-Rad) and then stripped and reprobed with the labeled 18S internal control
238 probe.

239

240 # III. Results and Discussion

241 **III-1. GC-biased characters of multiple repetitive ITS copies in the *H. sinensis***
242 **genome**

243 Generally, fungal genomes contain multiple repetitive copies of nrDNA ITS1-5.8S-ITS2
244 sequences. However, a single copy of the ITS1-5.8S-ITS2 sequences was identified in genome
245 assemblies ANOV01021709, LKHE01000582, and LWBQ01000008 of *H. sinensis* Strains Co18,
246 1229, and ZJB12195, respectively (Hu et al. 2013; Li et al. 2016a; Jin et al. 2020), probably because
247 the ITS repeats were discarded during the assembly of the genome shotgun sequences. Multiple
248 repetitive ITS copies were identified in 2 additional genome assemblies, JAAVMX000000000 and
249 NGJJ00000000, for *H. sinensis* Strains IOZ07 and CC1406-203, respectively (Liu et al. 2020; Shu et
250 al. 2020) (Table 1), where 17 ITS1-5.8S-ITS2 copies were found within JAAVMX000000000, and 7

full-length ITS1-5.8S-ITS2 copies and 2 partial ITS segments were found within NGJJ00000000.

Table 1. *GC contents of the repetitive ITS1-5.8S-ITS2 sequences from the genome assemblies of 2 H. sinensis strains and percent similarities of the repetitive ITS sequences to the sequences of GC-biased Genotype #1 (AB067721) and AT-biased Genotypes #4–6 and #15–17 of O. sinensis.*

Genome segment	Sequence range & direction	GC content	Percent similarity of the repetitive sequence	
			vs. Genotype #1 AB067721	*vs*. AT-biased Genotypes #4–6 and #15–17
JAAVMX010000002	18688917→18689407	64.8%	100%	86.7% – 89.9%
	18702095→18702586	64.5%	97.4%	82.8% – 88.0%
JAAVMX010000008	13823→14313	64.8%	100%	85.5% – 89.9%
	1199→1687	64.8%	99.2%	85.2% – 89.3%
JAAVMX010000017	9147←9637	64.8%	100%	85.5% – 89.9%
	21791←22281	64.8%	100%	85.5% – 89.9%
	34435←34925	64.8%	100%	85.5% – 89.9%
	47079←47569	64.8%	100%	85.5% – 89.9%
JAAVMX010000018	13381→13871	64.8%	100%	85.5% – 89.9%
	26076→26566	64.8%	100%	85.5% – 89.9%
	38771→39261	64.8%	100%	85.5% – 89.9%
	51467→51958	64.6%	98.8%	85.5% – 88.8%
	700→1186	64.5%	97.0%	82.6% – 87.6%
JAAVMX010000019	19404→19894	64.8%	100%	85.5% – 89.9%
	32048→32538	64.8%	100%	85.5% – 89.9%
	6233→6733	65.3%	94.5%	81.3% – 85.5%
	44729→45251	63.3%	90.8%	80.1% – 84.6%
NGJJ01000573	13133←13623	64.8%	99.4%	85.2% – 89.6%
	14624←15113	64.7%	99.4%	84.8% – 89.6%
	1←260 (ITS1 and partial 5.8S)	56.5%	98.5%	83.5% – 88.3%
NGJJ01000582	10703←11192	64.7%	99.8%	85.5% – 89.8%
	24068←24190 (only partial ITS2)	73.2%	98.1%	87.7% – 89.1%
NGJJ01000796	1514→2004	64.8%	100%	85.5% – 89.9%
NGJJ01000798	1177→1665	64.6%	99.6%	84.8% – 89.6%
NGJJ01000799	3295→3785	64.8%	100%	85.5% – 89.9%
	16666→17156	64.8%	100%	85.5% – 89.9%

Most of the repetitive ITS copies were highly homologous (>97%) to reference sequence AB067721 of GC-biased Genotype #1 *H. sinensis*. Two copies (6233→6733 and 44729→45251) within JAAVMX010000019 showed similarities of 94.5% and 90.8%, respectively, to AB067721 (Table 1). Two other copies, 18702095→18702586 within JAAVMX010000002 and 700→1186

261 within JAAVMX010000018, showed similarities of 97.4% and 97.0%, respectively (Table 1). All of
262 the repetitive ITS copies listed in Table 1 showed 80.1%–89.9% similarity to the sequences of AT-
263 biased Genotypes #4–6 and #15–17 of *O. sinensis* (Tables 1 and 2).

264 ***Table 2.*** *Percent similarities between nrDNA ITS1-5.8S-ITS2 sequences of 5 genome assemblies of*
265 *Genotype #1 H. sinensis and 17 O. sinensis genotype sequences. (Modified from Table 1 of Li* et al.
266 *2022d).*

GT	GenBank Accession #	% similarity between the *H. sinensis* genomic ITS segments and 17 *O. sinensis* genotypes				
		ANOV01021709 838→1504	LKHE01000582 2074→2740	LWBQ01000008 991739→992405	JAAVMX010000017 9029←9695	NGJJ01000799 3237→3903
#1	AB067721	99.7%	100%	99.6%	100%	100%
#2	MG770309	94.4%	94.7%	94.7%	94.7%	94.7%
#3	HM595984	95.9%	96.1%	95.6%	95.3%	95.3%
#7	AJ488254	94.7%	95.0%	95.3%	95.0%	95.0%
#8	GU246286	89.5%	89.5%	89.2%	89.5%	89.5%
#9	GU246288	95.0%	95.2%	94.6%	95.2%	95.2%
#10	GU246287	83.0%	83.2%	83.0%	83.2%	83.2%
#11	JQ695935	78.9%	79.1%	78.5%	79.1%	79.1%
#12	GU246296	94.3%	94.5%	93.9%	94.5%	94.5%
#13	KT339190	87.5%	87.5%	87.2%	87.5%	87.5%
#14	KT339178	89.3%	89.6%	89.1%	89.6%	89.6%
#4	AB067744	89.1%	89.4%	89.3%	89.4%	89.4%
#5	AB067740	86.3%	86.7%	86.2%	86.7%	86.7%
#6	KJ720572	85.5%	85.5%	85.5%	85.5%	85.5%
#15	KT232017	89.6%	89.9%	89.4%	89.9%	89.9%
#16	KT232019	87.4%	87.7%	87.6%	87.7%	87.7%
#17	KT232010	87.7%	88.1%	87.5%	88.1%	88.1%

267 *Note: "GT" represents genotype. GC-biased genotypes are listed in blue in the upper panel. AT-*
268 *biased genotypes are listed in red in the lower panel.*
269

270 These variable and slightly less variable GC-biased repetitive ITS copies contain predominantly
271 insertion/deletion mutant alleles, with much less predominant transversion mutant alleles and no or a
272 few transition mutant alleles (Table 3). The variable (94.5% and 90.8% similarity) ITS repeats,
273 6233→6733 and 44729→45251 within JAAVMX010000019, contained 22 and 32 insertion/deletion
274 alleles and 4 and 11 transversion alleles (Table 3; in blue in Figure 1); they contained only 1 and 5
275 transition mutant alleles, having allelic ratios of insertion/deletion/transversion *vs*. transition of 26:1
276 and 43:5, respectively. The slightly less variable (97.0–97.4% similarity) ITS repeats
277 18702095→18702586 within JAAVMX010000002 and 700→1186 within JAAVMX010000018

contained no transition alleles and predominantly included 13 and 15 insertion, deletion and transversion alleles, with allelic ratios of insertion/deletion/transversion *vs*. transition of 13:0 and 15:0 (Table 3; in green in Figure 1). Other repetitive ITS copies with 98.1%–99.8% similarity to AB067721 listed in Table 1 contained only a few insertion/deletion alleles.

Table 3. *Types of mutations in the variable and slightly less variable genomic repetitive ITS sequences compared with GC-biased Genotype #1 (AB067721).*

	Sequence range & direction	% similarity	Number of mutant alleles			Ins./Del./Transv. *vs*. Transit.
			Ins./Del.	Transv.	Transit.	
JAAVMX010000002	18702095→18702586	97.4%	13	0	0	13 : 0
JAAVMX010000018	700→1186	97.0%	10	5	0	15 : 0
JAAVMX010000019	6233→6733	94.5%	22	4	1	26 : 1
	44729→45251	90.8%	32	11	5	43 : 5

Note: "Ins./Del." indicates insertion/deletion point mutations. "Transv." denotes Transversion point mutations. "Transit." refers to transition point mutations.

All the repetitive ITS copies within genome assemblies JAAVMX000000000 and NGJJ00000000 listed in Table 1, for *H. sinensis* Strains IOZ07 and CC1406-203, and the single genomic ITS copies ANOV01021709, LKHE01000582, and LWBQ01000008 listed in Table 2, for Strains Co18, 1229, and ZJB12195, respectively (Hu et al. 2013; Jin et al. 2020; Li et al. 2016a; Liu et al. 2020; Shu et al. 2020), are GC-biased with GC contents of 64.7±2.3% (Table 1), similar to the GC content (64.8%) of the AB067721 Genotype #1 sequence. Regardless of the sequence similarity to the reference sequence AB067721 of Genotype #1, the repetitive ITS sequences contained predominantly insertion and deletion alleles and less frequent transversion alleles, while no or only a few transition alleles were found. These genetic features are distinct from AT-biased Genotypes #4–6 and #15–17 of *O. sinensis*, which included multiple scattered transition alleles (Figure 1) and contained 46.9%–55.2% AT bases and 44.8%–53.1% GC bases.

```
GT#1  AB067721  59        TCGAGTCACCACTCCC AAACCCCCTGCGAACACCACAGCAGTTGCCTCGGCGGGACCGCCCCGGCG CCCCAGGGCCCG 136
GT#3  FJ654148  1         CT----TG-------- ---A----C-T---------------------------------------- ------------ 83
JAAVMX010000019 19404     ---------------- ---------------------------------------------------- ------------ 19481
JAAVMX010000019 32048     ---------------- ---------------------------------------------------- ------------ 32125
JAAVMX010000002 18702095  ---------------- ------------------------ ---------------------------T------------ 18702172
JAAVMX010000018 700       --------- ------ ---------------------------------------T------------ ------------T 776
JAAVMX010000019 6233      ----------------C--------------------------------------------------- ------------ 6311
JAAVMX010000019 44729     ---------------- ---------------------------------------------------- ------------ 44806
GT#4  AB067744  55        ------T--------T --------------T--T--------------------------------- ---T-------- 132
GT#5  AB067740  59        ------T--T----TT ---------ATA--T--T-T-----------TA------------TA--- ---T-A------ 136
GT#6  KJ720572  11                         -------A----T--T-T-----------TA------------TA--- --TT-A------ 70
GT#15 KT232017  37        ------T--------T --------------T--T--------------------------------- ---T-------- 114
GT#16 KT232019  37        ------T--T-----T --------------T--T-T-----------TA------------TA--- ---T-A------ 114
GT#17 KT232010  37        ------T--T-----T --------------T--T-T-----------TA------------TA--- ---T-A------ 114

GT#1  AB067721  137       GACCAGG GCGCCC GCCGGAGGACCCCCAGAC CCTCCTGTCGCA GTGGCATCT CTCAGTCAAGAA GCAAGCAAAT 210
GT#3  FJ654148  84        ------- ------ ------------------ T----------- -----G--- ------------ ---------- 157
JAAVMX010000019 19482     ------- ------ ------------------ ------------ --------- ------------ ---------- 19555
JAAVMX010000019 32126     ------- ------ ------------------ ------------ --------- ------------ ---------- 32199
JAAVMX010000002 18702173  ------- ------A------------------ -------------A--------- ---- ------- ---------- 18702247
JAAVMX010000018 777       ------- ------ ------------------ ------------ --------- ---C-------- ---------- 850
JAAVMX010000019 6312      ------- ------ ------T-----------C------------ ---------C------------G---------- 6388
JAAVMX010000019 44807     -------C------ ------------------ ------------ --------- ------------ ---------- 44881
GT#4  AB067744  133       ---T--- ------ ------------TT---- ------A---T- --A------ --T---T----- ---------- 206
GT#5  AB067740  137       ---T-AA ------ ------------TT---- ------A-TAT- --A-T---- --T---T----- ---------- 213
GT#6  KJ720572  71        ---T--- ------ ------------TT---- ------A---T- --A-T---- --T---T----- ---------- 182
GT#15 KT232017  115       ---T--- ------ ------------TT---- ------A----- --A------ --T---T----- ---------- 188
GT#16 KT232019  115       ---T-AA ------ ------------TT---- ------A----- --A-T---- --T---T----- ---------- 188
GT#17 KT232010  115       ---T-AA ------ ------------TT---- ------A-TAT- --A-T---- --T---T----- ---------- 188

GT#1  AB067721  211       GAATCAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATC GATG AAGAACGCAGCGAAATGCGATAAGTAATGTGAA 288
GT#3  FJ654148  158       ---------------------------------------- ---- ---------------------------------- 235
JAAVMX010000019 19556     ---------------------------------------- ---- ---------------------------------- 19633
JAAVMX010000019 32200     ---------------------------------------- ---- ---------------------------------- 32277
JAAVMX010000002 18702248  -- ------------------------------------C---- ---------------------------------- 18702325
JAAVMX010000018 851       ---------------------------------------- ----A----------------------- --------- 928
JAAVMX010000019 6389      -------------- ---CAC-------------------- ---- --------------- --------- ------- 6463
JAAVMX010000019 44882     ---------------------------------------- ---- ---------------------------------- 44959
GT#4  AB067744  207       A---T--------T--T------------A---T-A----T A--A ----------------A-------------A-- 284
GT#5  AB067740  214       A---T----------T---------------A-----A----T A--A --A----------------A---------A-A-- 288
GT#6  KJ720572  183       A---T--------T---------------A-----A----- ---A -------T--T-----A-----------A-A-- 222
GT#15 KT232017  189       ----T--------T--T------------A---T-A----T A--A ----------------A-------------A-- 266
GT#16 KT232019  189       ----T--------T---------------A-----A----- ---A --A----------------A---------A-A-- 266
GT#17 KT232010  189       A---T--------T---------------A-----A----- ---A --A----------------A---------A-A-- 266

GT#1  AB067721  289       TTGCAGAATTCAGTGAACCA TCGAATCTTT GAACG CACATTG CGCCCGCCAGCACTCTGGCGGGCATGC  CTGTC 362
GT#3  FJ654148  236       -------------------- ---------- ----- ------- ---------------------------  ----- 309
JAAVMX010000019 19634     -------------------- ---------- ----- ------- ---------------------------  ----- 19707
JAAVMX010000019 32278     -------------------- ---------- ----- ------- ---------------------------  ----- 32351
JAAVMX010000002 18702326  -------------------- ----------T----- ----- - ------- ---------------------  ----- 18702398
JAAVMX010000018 929       ------------------- C---------- -----C-------C--------------------- -------  ----- 1000
JAAVMX010000019 6464      --------------------C---------- ----- -------C---------------------------  ----- 6539
JAAVMX010000019 44960     -------------------- ---------- ----- ------- ---------------------------CT----- 45035
GT#4  AB067744  285       ---T------T---A---T- ---------- A---- T-T-C-A -------T--------A----------  ----- 358
GT#5  AB067740  289       --A-------T---A---T- ---------- A---- --T---A -------T--------A----------  --A-- 362
GT#6  KJ720572  223       --AT------T---A----- ---------- A---- T-T---A -------T--------A----------  ----- 296
GT#15 KT232017  267       ---T------T---A---T- ---------- A---- T-T---A -------T--------A----------  ----- 340
GT#16 KT232019  267       --A-------T---A---T- ---------- A---- T-T---A -------T--------A----------  ----- 340
GT#17 KT232010  267       --A-------T---A---T- ---------- A---- --T---- -------T--------A----------  --A-- 340

GT#1  AB067721  363       CGAGCGTCATCTCAACCCTCGAG CCCCCCGCCTCGCGG  CGGCGGGGCCCGGCCTT GGGGGTCACGGCC CCGCGCC 437
GT#3  FJ654148  310       ------------------------ --------------  ----------------- ------------- ------- 384
JAAVMX010000019 19708     ------------------------  --------------  ----------------- ------------- ------- 19781
JAAVMX010000019 32352     ------------------------  --------------  ----------------- ------------- ------- 32425
JAAVMX010000002 18702399  ------------------------ ----- ---------  ----------------- ------------- ------- 18702472
JAAVMX010000018 1001      ------------------------ ------ --------  -----------------G------------- ---- -- 1074
JAAVMX010000019 6540      ------------------------  --------------  ------------- ---- -------------C------- 6613
JAAVMX010000019 45036     ------------------------C----------A----TA--CG---C--------- ------TC-----AA------ 45119
GT#4  AB067744  359       -------T-C--T------TA-- ---------------  ----------------- A------------ --A---- 432
GT#5  AB067740  363       -------T----T---------- ---------------  ----------T------- A-----T------ ------- 436
GT#6  KJ720572  297       -------T-C--T----                                                                313
GT#15 KT232017  341       -------T----T---------- ---------------  ----------T------- A-----T------ ------- 414
GT#16 KT232019  341       -------T-C--T------TA-- ---------------  ----------------- A------------ --A---- 414
GT#17 KT232010  341       -------T----T---------- ---------------  ----------T------- A-----T------ ------- 414

GT#1  AB067721  438       G   CCCCCTAAACGC AG       TGGCGACC  CCGCCGCGGCTCC CCTGCGCAGTAGCT  CG  CTGAGAACCT 499
GT#3  FJ654148  385       -   ---------T-- --       --------  ------------- --------------  --  -C---G---C 446
JAAVMX010000019 19782     -   ------------ --       --------  ------------- --------------  --  ---------- 19844
JAAVMX010000019 32426     -   ------------ --       --------  ------------- --------------  --  ---------- 32488
JAAVMX010000002 18702473  -   ------------ --       --------  ------------- --------------  --  ---------- 18702534
JAAVMX010000018 1075      -   ------------ --       --------  ------------- --------------  --  ---------- 1136
JAAVMX010000019 6614      -    -----------G--G      --------GC------------- --------------  --GCT--------- 6680
JAAVMX010000019 45120     -TCC-T---GT----- --CCTAAGTG-------  ----GA-------A----T-A-------CA--  ---------- 45189
GT#4  AB067744  433       -   ------------ --       -A-T----  ------------- ---A----A-----  TA  --A------- 495
GT#5  AB067740  437       -   ----T------T --       -A------  ------------- ---A--TG------  -A  -C-------- 529
GT#6  KJ720572  313
GT#15 KT232017  415       -   ------------T --       -A------  ------------- ---A--T-A-----  -A  -C-------- 477
GT#16 KT232019  415       -   ------------ --       -A------  ------------- ---A----A-----  TA  --A------- 477
GT#17 KT232010  415       -   -----------T --       -A------  ------------- ---A--T-A-----  -A  -C-------- 477

GT#1  AB067721  500        CGCACCGGGAGCGCGGA     GGCGG TCACGCC  TGG    AA ACCACC   ACACCCTCCA              549
GT#3  FJ654148  447        --------A-------     ----- -------  ---    -- ------   ---     ---             492
JAAVMX010000019 19845      ------------------     ----- ----- -  ---    -- ------   ----------             19894
JAAVMX010000019 32489      ------------------     ----- ----- -  ---    -- ------   ----------             32537
JAAVMX010000002 18702535   ------------------     ----- -------GTGT-    -- ------   ----------             18702586
JAAVMX010000018 1137       ------------------     ----- ----- -  ---    -- ------   ----------             1186
JAAVMX010000019 6681      C------------------     ----- ----- -  ---    -- ------ACA----------             6733
JAAVMX010000019 45190      ------------------CCCAT-----C-------  ---GATG--C------ACC----------             45251
GT#4  AB067744  496        -----TA----------     ----- -------  --A    -- -----T   -T------              543
GT#5  AB067740  530        --------A--------     ----- -------  --A    -- -----T   --G--                 544
GT#6  KJ720572  313
GT#15 KT232017  478        --------A--------     ----- -------  --A    -- -----T   --G--                 522
GT#16 KT232019  478        -----TA----------     ----- -------  --A    -- -----T   -T------              525
GT#17 KT232010  478        --------A--------     ----- -------  --A    -- -----T   --G--                 522
```

298

Figure 1. ITS sequence alignment of AB067721, repetitive ITS copies within genome JAAVMX000000000 of H. sinensis Strain IOZ07, and AT-biased genotypes of O. sinensis.

The ITS sequences contain complete or partial ITS1-5.8S-ITS2 nrDNA segments. "GT" denotes the O. sinensis genotype. The underlined sequence in black represents the 5.8S gene of GC-biased Genotype #1 H. sinensis. AB067721 is the ITS sequence of GC-biased Genotype #1 H. sinensis. Genome assemblies JAAVMX010000002, JAAVMX010000018 and JAAVMX010000019 were obtained from H. sinensis Strain IOZ07 (Shu et al. *2020). One copy each within JAAVMX010000002 and JAAVMX010000018 indicated in* ***green*** *shows 97.4% or 97.0% similarity to AB067721. JAAVMX010000019 contains 4 repetitive ITS copies, including 2 copies in* ***black*** *(19404→19894 and 32048→32537), which are 100% identical to AB067721, and 2 other copies in* ***blue*** *(6233→6733 and 44729→45251), with 94.5% and 90.8% similarities to AB067721. Sequences in* ***red*** *(AB067744, AB067740, KJ720572, KT232017, KT232019, and KT232010) represent AT-biased Genotypes #4–6 and #15–17 of O. sinensis, respectively. The underlined "GAATTC" sequences in* ***purple*** *represent the Eco*RI *endonuclease cleavage sites in the sequences of GC-biased Genotype #1 and GC-biased genome assembly JAAVMX000000000. Eco*RI *endonuclease cleavage sites are absent in AT-biased ITS sequences due to a single-base C-to-T (cytosine→threonine) transition. Hyphens indicate identical bases, and spaces denote unmatched sequence gaps.*

Selker and his colleague (1985, 1990, 1997, 2002) reported frequent C-to-T (cytosine→thymine) transition mutations and concomitant epigenetic methylation of nearly all remaining cytosine residues in fungal 5S RNA genes associated with transcriptional silencing. Gladyshev (2017) summarized that RIP (1) "occurs specifically in haploid parental nuclei that continue to divide by mitosis in preparation for karyogamy and ensuing meiosis;" (2) "detects duplications of chromosomal DNA above a certain length threshold (~0.4 kbp);" (3) "mutates cytosines on both strands of each DNA duplex in a pairwise fashion (e.g., preferentially affecting none or both repeated sequences);" (4) "occasionally spreads from duplicated sequences into neighboring nonrepetitive regions;" and (5) "remains unaffected in *mei-2* crosses that exhibit a strong defect in meiotic pairing of homologous chromosomes."

The variable repetitive ITS copies characterized by multiple scattered insertion, deletion and transversion point mutations (Table 3; Figure 1) might not have been generated through RIP mutagenesis, which theoretically causes C-to-T (cytosine→threonine) and G-to-A (guanine→adenine) transition mutations. Thus, the hypothesis of Li et al. (2020b) that RIP mutagenesis generated mutant AT-biased genotypes of *O. sinensis* within the single genome of GC-biased Genotype #1 *H. sinensis* lacks supporting evidence.

III-2. Genomic independence of multiple genotypes of *O. sinensis*

Answering the question of whether the AT-biased sequences of *O. sinensis* are components of a single genome of Genotype #1 *H. sinensis* will provide key evidence for the "ITS pseudogene"

334 hypothesis. However, the sequences of mutant Genotypes #2–17 of *O. sinensis* were not present in
335 genome assemblies ANOV00000000, JAAVMX000000000, LKHE00000000, LWBQ00000000 and
336 NGJJ00000000 of *H. sinensis* strains Co18, IOZ07, 1229, ZJB12195, and CC1406-203, respectively,
337 as shown in Tables 1–2 and Figure 1 (Hu et al. 2013; Zhu & Wu 2015; Li et al. 2016a, 2016b, 2020a,
338 2022a, 2022d, 2022e; Zhu & Li 2017; Zhu et al. 2018, 2019; Jin et al. 2020; Liu et al. 2020; Shu et
339 al. 2020). The GC-biased sequences of repetitive ITS copies in the genomes of *H. sinensis* strains are
340 80.1%–89.9% similar to the sequences of AT-biased Genotypes #4–6 and #15–17 of *O. sinensis*
341 (Table 1). The sequences of Genotypes #2–17 (both GC- and AT-biases) of *O. sinensis* are absent in
342 the genome of Genotype #1 *H. sinensis,* as shown in Table 2, but instead belong to the genomes of
343 independent *O. sinensis* fungi (Stensrud et al. 2008; Xiao et al. 2009; Zhu and Wu 2015; Li et al.
344 2016b; 2022d, 2022e; Zhu and Li 2017). The nonresidency measure of AT-biased *O. sinensis*
345 genotypes does not support the notion that the AT-biased sequences are "ITS pseudogene"
346 components of "a single genome" of GC-biased Genotype #1 *H. sinensis*.

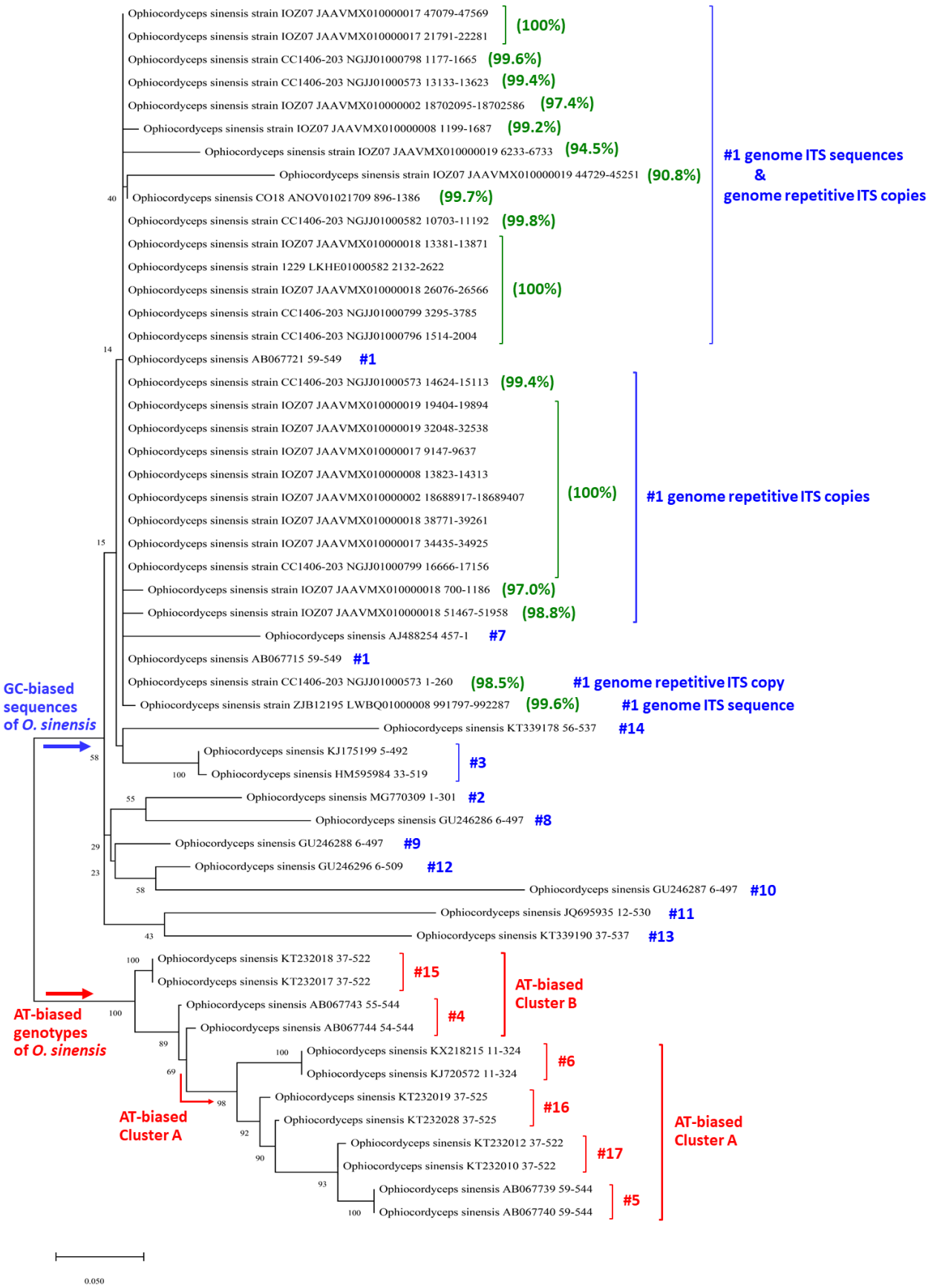
347 The sequences of AT-biased Genotypes #4–6 and #15–17 of *O. sinensis* shown in red in Figure
348 1 contain numerous scattered transition alleles. They were generated through long-term accelerated
349 nrDNA evolution from a common ancestor during a long history of evolution, regardless of whether
350 this occurred *via* RIP or other mutagenic mechanisms. However, the multiple scattered transition
351 point mutations in the AT-biased genotypes are genetically and phylogenetically distinct from the
352 predominant insertion, deletion and transversion point mutations, as observed in the sequences of
353 repetitive ITS copies within the genome of GC-biased Genotype #1 *H. sinensis* (Figures 1–2 and
354 Tables 1–3). The notion that RIP mutagenesis induces the formation of AT-biased repetitive ITS
355 copies in the single genome of Genotype #1 *H. sinensis* lacks genomic evidence. Even if it is proven
356 that AT-biased genotypes were indeed induced by RIP mutagenesis, there has been no proof that they
357 are genomic components of the GC-biased Genotype #1 genome that emerged after a new GC-biased
358 Genotype #1 "ITS haplotype was generated" as "ITS pseudogenes … in a single genome", as assumed
359 by Li et al. (2013, 2020b).

360 GC-biased genotypes, including the GC-biased genomic repetitive ITS copies, might have
361 evolved in parallel with AT-biased genotypes from an ancient common ancestor during long-term
362 evolution, as assumed by Stensrud et al. (2007). Zhang et al. (2009) detected GC-biased Genotype
363 #3 in natural *C. sinensis* specimens collected from the Nyingchi District of Tibet and hypothesized
364 that "Nyingchi might have acted as **a center of origin** of *O. sinensis* and *O. sinensis* might have been

365 transmitted from Nyingchi to other locations". Accordingly, GC-biased Genotype #3 might represent
366 a modern immediate ancestor of Genotype #1 *H. sinensis* that has been identified in other production
367 locations during long-term evolution. The authors described an evolution path from GC-biased
368 Genotype #3 to Genotype #1 and transmission of *O. sinensis* from the "center of origin … spread
369 eastward and westward first among southern regions on the Tibetan Plateau" and then "spread from
370 south to north … at a later time". However, this evolutionary path and transmission of *O. sinensis* are
371 not evolutionarily related to AT-biased genotypes.

372 The genomic repetitive ITS copies, regardless of their similarities to Genotype #1 AB067721,
373 were phylogenetically clustered into the GC-biased clades in blue alongside the Bayesian tree in
374 Figure 2. Apparently, they were phylogenetically distant from the AT-biased clades containing
375 Genotypes #4–6 and #15–17 of *O. sinensis* shown in red alongside in the tree. Stensrud et al. (2007)
376 concluded that the variation in the 5.8S gene sequences of AT-biased genotypes "far exceeds what is
377 normally observed in fungi … even at higher taxonomic levels (genera and family)". Xiao et al. (2009)
378 concluded that the sequences of AT-biased genotypes likely belong to independent *O. sinensis* fungi.

379 Figure 2 includes the 4 genomic repetitive ITS copies with low or relatively low similarities
380 shown in Table 3, which are individually scattered in the GC-biased clades. The 2 low-similarity
381 genomic repetitive ITS copies, 6233→6733 (94.5%) and 44729→45251 (90.8%), within genome
382 assembly JAAVMX010000019 (*cf*. Tables 1 and 3), showed greater phylogenetic (horizontal lines)
383 distances from the GC-biased clades. These results indicate that genetic variations in the repetitive
384 ITS sequences shown in blue in Figure 1 had an impact on Bayesian phylogenetic clustering (Figure
385 2).


Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 47079-47569
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 21791-22281
(100%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000798 1177-1665 (99.6%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000573 13133-13623 (99.4%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000002 18702095-18702586 (97.4%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000008 1199-1687 (99.2%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 6233-6733 (94.5%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 44729-45251 (90.8%)
40
Ophiocordyceps sinensis CO18 ANOV01021709 896-1386 (99.7%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000582 10703-11192 (99.8%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 13381-13871
Ophiocordyceps sinensis strain 1229 LKHE01000582 2132-2622
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 26076-26566
(100%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000799 3295-3785
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000796 1514-2004
#1 genome ITS sequences
&
genome repetitive ITS copies
14
Ophiocordyceps sinensis AB067721 59-549 #1
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000573 14624-15113 (99.4%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 19404-19894
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 32048-32538
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 9147-9637
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000008 13823-14313
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000002 18688917-18689407
(100%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 38771-39261
15
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 34435-34925
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000799 16666-17156
#1 genome repetitive ITS copies
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 700-1186 (97.0%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 51467-51958 (98.8%)
Ophiocordyceps sinensis AJ488254 457-1 #7
Ophiocordyceps sinensis AB067715 59-549 #1
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000573 1-260 (98.5%) #1 genome repetitive ITS copy
Ophiocordyceps sinensis strain ZJB12195 LWBQ01000008 991797-992287 (99.6%) #1 genome ITS sequence
GC-biased sequences of O. sinensis
Ophiocordyceps sinensis KT339178 56-537 #14
58
Ophiocordyceps sinensis KJ175199 5-492
100
Ophiocordyceps sinensis HM595984 33-519
#3
55
Ophiocordyceps sinensis MG770309 1-301 #2
Ophiocordyceps sinensis GU246286 6-497 #8
29
Ophiocordyceps sinensis GU246288 6-497 #9
23
Ophiocordyceps sinensis GU246296 6-509 #12
58
Ophiocordyceps sinensis GU246287 6-497 #10
Ophiocordyceps sinensis JQ695935 12-530 #11
43
Ophiocordyceps sinensis KT339190 37-537 #13
100
Ophiocordyceps sinensis KT232018 37-522
Ophiocordyceps sinensis KT232017 37-522
#15
AT-biased Cluster B
AT-biased genotypes of O. sinensis
100
Ophiocordyceps sinensis AB067743 55-544
Ophiocordyceps sinensis AB067744 54-544
#4
89
100
Ophiocordyceps sinensis KX218215 11-324
Ophiocordyceps sinensis KJ720572 11-324
#6
69
98
Ophiocordyceps sinensis KT232019 37-525
Ophiocordyceps sinensis KT232028 37-525
#16
AT-biased Cluster A
92
90
Ophiocordyceps sinensis KT232012 37-522
Ophiocordyceps sinensis KT232010 37-522
#17
93
Ophiocordyceps sinensis AB067739 59-544
100
Ophiocordyceps sinensis AB067740 59-544
#5
AT-biased Cluster A
0.050

386

Figure 2. A Bayesian majority rule consensus phylogenetic tree.
Twenty-eight ITS segments of the genome assemblies (ANOV01021709, LKHE01000582, LWBQ01000008, JAAVMX010000002, JAAVMX010000008, JAAVMX0100000017, JAAVMX0100000018, JAAVMX010000019, NGJJ01000573, NGJJ01000582, NGJJ01000796, NGJJ01000798, and NGJJ01000799) of H. sinensis strains (Co18, 1229, ZJB12195, IOZ07, and CC1406-203, respectively) and 25 ITS sequences of GC-biased Genotypes #1–3 and #7–14 (in blue alongside the tree) and AT-biased Genotypes #4–6 and #15–17 of O. sinensis (in red alongside the tree) were analyzed phylogenetically using MrBayes v3.2.7 software (cf. Section II-4). Genome assemblies JAAVMX000000000 and NGJJ00000000 contain multiple repetitive ITS copies (cf. Table 1). The percent similarities of the genomic ITS sequences with the representative Genotype #1 sequence (AB067721) are shown in green alongside the tree.

Li et al. (2020b) hypothesized that (1) AT-biased *O. sinensis* genotypes "could have emerged either before or after" a new GC-biased "ITS haplotype was generated" and that (2) RIP mutagenesis is "responsible for nrDNA pseudogene formation (*i.e.*, AT-biased genotypes)". These 2 assumptions support the "ITS pseudogene" hypothesis that AT-biased sequences coexist "in a single genome" of GC-biased Genotype #1, as repeatedly stated by Li et al. (2013; 2017, 2019, 2020b). However, the genomic residency measure results obtained from different perspectives confirmed that the sequences of AT-biased genotypes are not genomic components of GC-biased Genotype #1 but instead belong to independent genomes of *O. sinensis* (further discussion below). AT-biased genotypes possessing numerous scattered transition alleles might have independently evolved through RIP or other mutagenic and evolutionary mechanisms, in parallel with GC-biased genotypes, during a long history of evolution from the common genetic ancestor and coexist in different combinations in different compartments of natural *C. sinensis* as independent *O. sinensis* fungal species, even "at higher taxonomic levels (genera and family)" (Stensrud et al. 2007; Xiao et al. 2009).

III-3. Inconsistency of the genetically divergent Genotype #1 *H. sinensis* strains

Li et al. (2013) reported the identification of heterogeneous ITS sequences of both GC-biased Genotype #1 and AT-biased Genotype #5 in 8 of 15 clones after 25 days of inoculation with natural *C. sinensis* mono-ascospores and hypothesized that Genotypes #1 and #5 coexisted "in a single genome". This speculation is irreconcilable with the simultaneous identification of 7 other clones derived from the same mono-ascospores that contained only homogenous sequences of Genotype #1. It is even more inexplicable why Li et al. (2013, 2017, 2019, 2020b) repeatedly called all AT-biased genotypes divergent pseudogenes, which coexisted "in single individuals of *O. sinensis*" or "in a single genome" of GC-biased Genotype #1 (the functional counterpart), while the researchers did not

420 identify AT-biased Genotypes #4, #6, and #15–17 in all 15 clones obtained from the *C. sinensis* mono-
421 ascospores.

422 Bushley et al. (2013), who were collaborators of Li et al. (2013), reported that *C. sinensis*
423 ascospores are multicellular heterokaryons with mono-, bi-, and trinucleated cellular structures.
424 Zhang and Zhang (2015) wondered whether the binucleated cells (also trinucleated cells) of *C.*
425 *sinensis* hyphae and ascospores contain homogenous or heterogeneous genetic material, meaning that
426 these polykaryotic cells may contain heterogeneous sets of chromosomes and genomes. Furthermore,
427 various mono-nucleated hyphal and ascosporic cells of *C. sinensis* might also contain heterogeneous
428 genetic material.

429 By aggregating the information from molecular and genomic studies (Hu et al. 2013; Li et al.
430 2013, 2016a; Jin et al. 2020; Liu et al. 2020; Shu et al. 2020), different genomic DNA samples of
431 "pure" *H. sinensis* strains were examined and can be grouped as follows.

432 (1) Group 1 of *H. sinensis* strains, consisting of pure, homokaryotic anamorphic *H. sinensis* strains
433 (1229, CC1406-203, Co18, IOZ07, and ZJB12195) (Hu et al. 2013; Li et al. 2016a; Jin et al.
434 2020; Liu et al. 2020; Shu et al. 2020). Total genomic DNA was isolated from these strains and
435 subjected to genome-wide sequencing. The genomes contained no AT-biased genotype sequences.

436 (2) Group 2 of *H. sinensis* strains, consisting of 7 clones (Strains 1207, 1218, 1219, 1221, 1222, 1225
437 and 1229) among 15 clones derived from 25 days of incubation of natural *C. sinensis* mono-
438 ascospores (Li et al. 2013). Total genomic DNA was isolated from these strains and shown to
439 contain the homogenous ITS sequence of Genotype #1 *H. sinensis* but included no AT-biased
440 sequences of *O. sinensis* genotypes.

441 (3) Group 3 of "*H. sinensis*" strains, including 8 other clones (Strains 1206, 1208, 1209, 1214, 1220,
442 1224, 1227 and 1228) obtained from the incubation of *C. sinensis* mono-ascospores for 25 days
443 (Li et al. 2013). Total genomic DNA was isolated from these strains and showed genetic
444 heterogeneity, with the coexistence of GC-biased Genotype #1 and AT-biased Genotype #5 of *O.*
445 *sinensis*.

446 The *H. sinensis* strains of Groups 1 and 2 are similar, possessing a homogenous GC-biased
447 Genotype #1 genome, although genome assemblies JAAVMX000000000 and NGJJ00000000 of *H.*
448 *sinensis* Strains IOZ07 and CC1406-203 contained multiple GC-biased repetitive ITS copies (*cf.*
449 Tables 1 and 3 and Figures 1–2) (Hu et al. 2013; Li et al. 2013, 2016a; Jin et al. 2020; Liu et al. 2020;
450 Shu et al. 2020). Although a single copy of ITS sequences was identified in genome assemblies

451 ANOV00000000, LKHE00000000, and LWBQ00000000 of *H. sinensis* Strains Co18, 1229, and
452 ZJB12195, respectively (Hu et al. 2013; Li et al. 2016a; Jin et al. 2020), the genomes of these strains
453 may theoretically contain more repetitive ITS copies that could have been discarded during the
454 assembly of genome shotgun sequences. Conceivably, reassembling the genome shotgun reads may
455 very likely identify additional repetitive ITS copies in the genomes of Group 1 Strains Co18, 1229,
456 and ZJB12195, among which Strain 1229 is also a Group-2 strain derived from mono-ascospores by
457 Li et al. (2013, 2016a). Similarly, the genome sequencing of Group-2 strains may very likely identify
458 multiple repetitive ITS copies.

459 Although the anamorphic *H. sinensis* strains in Groups 2 and 3 were derived after 25 days of
460 liquid culture of the same *C. sinensis* mono-ascospores (Li et al. 2013), they are genetically different
461 and may likely have multicellular heterokaryotic mycelial structures, similar to the mycelia derived
462 from microcycle conidiation of the *H. sinensis* ascospores initially cultured on the solid medium plate
463 for 30 days and subsequently in liquid medium for 10–53 days, as shown in Figures 2–3 of (Xiao et
464 al. 2011). Li et al. (2013) reported that the Group 3 strains included cocultured GC-biased Genotype
465 #1 and AT-biased Genotype #5 of *O. sinensis* fungi, which were genomically independent and
466 coexisted in heterokaryotic mycelial cells, indicating either diploid/polyploid mycelial cells
467 containing heterogeneous sets of chromosomes/genomes or impure mycelia, each containing a
468 different chromosome/genome. However, the authors did not identify any haploid clones containing
469 only AT-biased Genotype #5, indicating that this AT-biased *O. sinensis* fungus may not exhibit
470 independent culturability under the *in vitro* experimental settings, although other studies have
471 reported the identification of single AT-biased Genotype #4 or #5 of *O. sinensis*, without coculture
472 with GC-biased Genotype #1, in cultures of natural and cultivated *C. sinensis* under different
473 experimental settings (Mao et al. 2013; Wei et al. 2016). Unfortunately, Li et al. (2013, 2017, 2019,
474 2020b) neglected the differences in genetic material between Group 2 and 3 strains, which were
475 presumably derived from the same or different cells of multicellular heterokaryotic ascospores
476 (Bushley et al. 2013; Zhang and Zhang 2015; Li et al. 2020a, 2022d, 2022e).

477 In addition to the genome assembly LKHE00000000 derived from the Group-2 *H. sinensis* Strain
478 1229 by Li et al. (2016a), the genome sequencing of one of the heterogeneous Group-3 strains may
479 provide critical genomic evidence to validate the notion that AT-biased Genotype #5 of *O. sinensis*
480 coexists as an "ITS pseudogene … in a single genome" of GC-biased Genotype #1. Unfortunately,
481 the authors elected not to do so since 2012 after obtaining the Group-3 strains, or perhaps failed in

482 such an endeavor due to the technical difficulties in shotgun sequence assembly because of the
483 heterogeneous sequences from more than one set of genomes. Consequently, they elected willingly
484 or unwillingly to pass on the opportunity to experimentally confirm their "ITS pseudogene"
485 hypothesis that AT-biased Genotype #5 and GC-biased Genotype #1 coexist "in a single genome".
486 Studies from other groups that are reviewed in this paper have confirmed that the sequences of all
487 AT-biased genotypes belong to the genomes of independent fungi (Stensrud et al. 2007; Xiao et al.
488 2009; Zhu et al. 2010; Yao et al. 2011; Li et al. 2016, 2020a, 2022a, 2022d, 2022e) and that
489 heterokaryotic ascospores and hyphae with multiple mono-, bi-, and tri-nucleated cells contain
490 heterogeneous sets of chromosomes and genomes (Bushley et al. 2013; Zhang and Zhang 2015).
491 Li et al. (2013) did not report the identification of ITS sequences of AT-biased Genotypes #4, #6,
492 and #15–17 of *O. sinensis* in genomic DNA preparations that were extracted from the 15 ascosporic
493 strains in Groups 2 and 3. The expansion of the "ITS pseudogene" hypothesis from only Genotype
494 #5 to all AT-biased genotypes was apparently an overinterpretation with no supporting evidence. As
495 shown in Tables 1–3 and Figures 1–2, genomic and phylogenetic analyses did not support the notion
496 that the "diverged ITS paralogs" (*i.e.*, AT-biased Genotypes #4–6 and #15–17) coexist "in a single
497 genome" of *O. sinensis*. Li et al. (2022d) demonstrated that the sequences of Genotypes #4 and #15
498 of AT-biased Cluster-B (*cf*. Figure 2) were absent in the genomic DNA pool extracted from natural *C.*
499 *sinensis* ascospores. Instead, sequences of these genotypes belong to independent genomes of *O.*
500 *sinensis* fungi, regardless of whether the AT-biased genotypes might have been generated *via* RIP or
501 other mutagenic mechanisms during the long evolutionary history of *O. sinensis*.
502

503 III-4. Differential co-occurrence and dynamic alteration of multiple *O. sinensis* 504 genotypes in the compartments of natural *C. sinensis* during maturation

505 A BLAST search of the GenBank database revealed 668 ITS sequences of *O. sinensis* (GenBank
506 taxid: 72228), 228 of which belong to AT-biased Genotypes #4–6 and #15–17 (Li et al. 2016b, 2022d,
507 2022e; Zhu et al. 2018, 2019; Zhu & Li 2017). After Kinjo and Zang (2001) reported for the first time
508 the 3 phylogenetic clades of mutant *O. sinensis* sequences, Stensrud et al. (2007) and Zhang et al.
509 (2013) analyzed the phylogeny of 71 and 397 ITS sequences, respectively, and confirmed the 3
510 phylogenetic clades, *i.e.*, Groups A–C (GC-biased Genotype #1 and AT-biased Genotypes 4–5),
511 which consisted of correct *O. sinensis* sequences. Additional GC- and AT-biased *O. sinensis*
512 genotypes have been successively discovered (*cf*. Table 2) (Yang et al. 2008; Xiao et al. 2009; Zhang

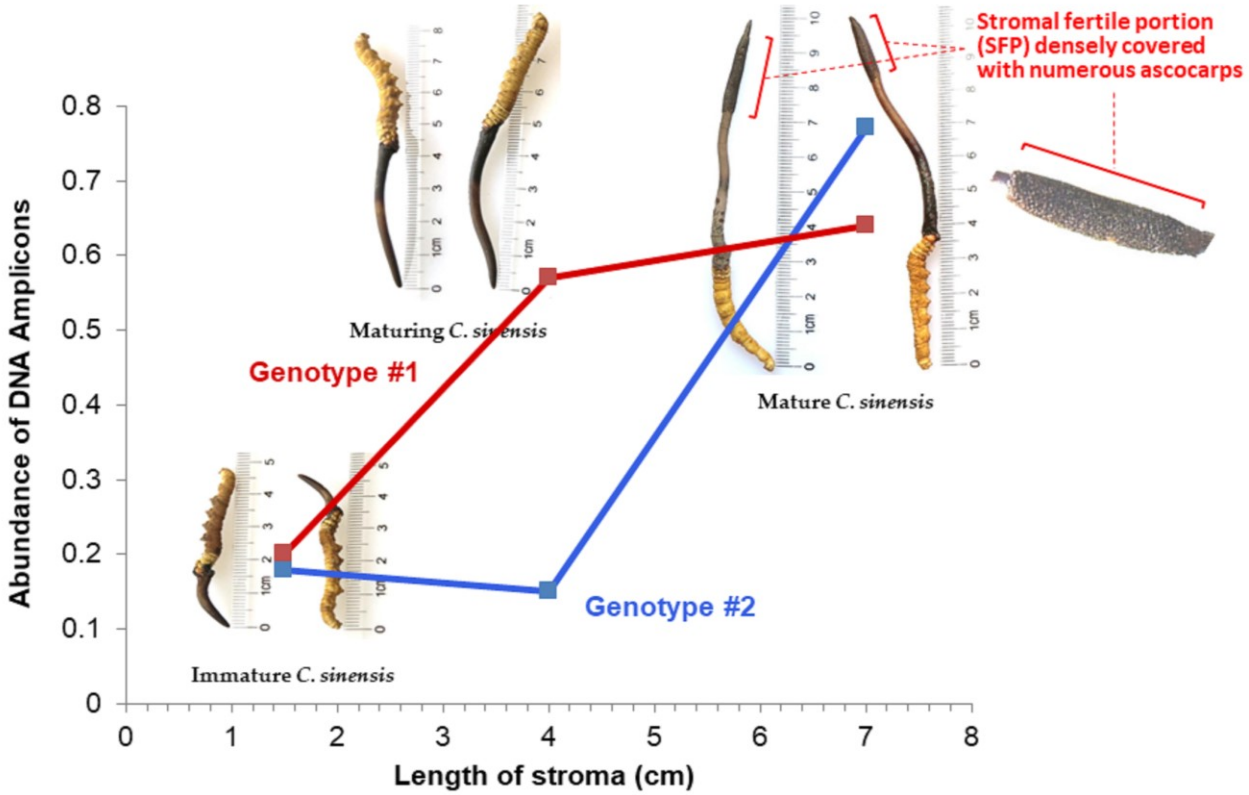
513 et al. 2009; Zhu et al. 2010, 2012; Chen et al. 2011; Zhu and Wu 2015; Li et al. 2013, 2016, 2022a,
514 2022d,2022e; Zhu and Li 2017). Molecular mycology studies confirmed the genetic heterogeneity of
515 natural *C. sinensis* with differential co-occurrence of several GC- and AT-biased genotypes of *O.*
516 *sinensis* in different combinations in the different compartments of natural *C. sinensis,* and the
517 biomasses/abundances of the genotypes were altered dynamically during *C. sinensis* maturation
518 (Xiao et al. 2009; Zhu et al. 2010, 2012, 2018, 2019; Gao et al. 2011, 2012; Yao et al. 2011; Mao et
519 al. 2013; Li et al. 2016b, 2016c, 2022a, 2022c, 2022d, 2022e).
520 Many studies have used pairs of "universal" primers (*e.g.*, *ITS1–ITS5*) for PCR and identified one
521 of the GC- and AT-biased genotypes of *O. sinensis* in natural *C. sinensis* specimens collected from
522 different production areas or cultivated *C. sinensis* (Engh 1999; Zhao et al. 1999; Kinjo and Zang
523 2001; Liu et al. 2001; Stensrud et al. 2005; Wei et al. 2006, 2016; Chen et al. 2011; Mao et al. 2013).
524 Wei et al. (2016) reported an industrial artificial cultivation project and identified a single
525 teleomorphic AT-biased Genotype #4 of *O. sinensis* in cultivated *C. sinensis* and a single
526 teleomorphic GC-biased Genotype #1 in natural *C. sinensis* as the teleomorph control, indicating the
527 existence of at least 2 *O. sinensis* teleomorphs. This study revealed a species contradiction between
528 the anamorphic inoculants (reportedly 3 *H. sinensis* strains) and teleomorphic AT-biased Genotype
529 #4 in cultivated *C. sinensis*.
530 Other studies optimized and combined the use of a variety of molecular techniques and identified
531 naturally cooccurring GC- and AT-biased genotypes of *O. sinensis* in different combinations in the
532 caterpillar body, stroma, stromal fertile portion with numerous ascocarps, and ascospores of natural
533 *C. sinensis* (Yang et al. 2008; Xiao et al. 2009; Zhu et al. 2010, 2012, 2018, 2019; Gao et al. 2011,
534 2012; Yao et al. 2011; Zhu and Wu 2015; Li et al. 2016b, 2020a, 2022a, 2022c, 2022d, 2022e; Zhu
535 and Li 2017). Xiao et al. (2009) used genotype-specific primers for the first time and a nested PCR
536 approach to amplify mutant sequences of *O. sinensis* from the stroma and caterpillar body of
537 immature and mature *C. sinensis* specimens collected from geographically remote production areas.
538 They reported the consistent occurrence of GC-biased Genotype #1 and inconsistent occurrence of
539 AT-biased Genotypes #4–5 of *O. sinensis* and then concluded that the AT-biased genotypes likely
540 belonged to independent fungal species. Two of the authors also identified the first sequence
541 (EU555436) of AT-biased Genotype #6 and uploaded it to GenBank. The use of genotype-specific
542 primers combined with touchdown PCR, restriction fragment length polymorphism (RFLP), cloning-
543 based amplicon sequencing, SNP mass spectrometry genotyping techniques, etc., in other studies

544 greatly improved the detection specificity and efficiency and facilitated the identification of
545 differential coexistence and dynamic alterations of GC- and AT-biased genotypes of *O. sinensis* in
546 the caterpillar body, stroma, stromal fertile portion with numerous ascocarps, and ascospores of
547 natural *C. sinensis* during maturation (Zhu et al. 2010, 2012, 2019; Chen et al. 2011; Gao et al. 2011,
548 2012; Yao et al. 2011; Zhu and Wu 2015; Li et al. 2016b, 2020a, 2022a, 2022c, 2022d, 2022e; Zhu
549 and Li 2017). To date, 17 genotypes of *O. sinensis* have been identified, which are genetically and
550 phylogenetically distinct (Table 2 and Figures 1–2).

551 In contrast to the inconsistent cooccurrence of Genotypes #1 and #4–5 (Xiao et al. 2009), Zhu et
552 al. (2010) used *H. sinensis*-specific primers, *Hsprp1*/*Hsprp3*, and identified the consistent coexistence
553 of GC-biased Genotypes #1 and #2 in the stromata of immature, maturing, and mature *C. sinensis*
554 specimens collected in Kangding County of Sichuan Province, China. Notably, the ITS sequences of
555 the GC-biased genotypes contained an *Eco*RI endonuclease cleavage site (G▾AATTC) from
556 nucleotides 294–299 of AB067721 (a representative sequence of Genotype #1 *H. sinensis*), which is
557 underlined in purple in Figure 1 and was lost in all AT-biased genotypes due to a single base C-to-T
558 (cytosine→thymine) transition mutation (GAATTT). Sequencing the target amplicons (440 bp, based
559 on the design of primers) obtained ITS sequences of Genotype #1 *H. sinensis* as shown in Figure 10
560 of (Zhu et al. 2010). A faster migrating amplicon (320 bp) on the agarose gel showed Genotype #2
561 ITS sequences (MG770309).

562 Due to the possibility that the predominant 440-bp amplicons might not be homogeneous based
563 on the various similarities of primers *Hsprp1*/*Hsprp3* to the ITS sequences of various *O. sinensis*
564 genotypes (although only a single sequence was obtained from the target amplicons), the 440-bp
565 amplicons obtained from immature, maturing and mature stromata were subjected to overnight *Eco*RI
566 digestion. Two digested fragments (240 and 200 bp) were generated with extremely low densities in
567 the immature stroma, representing GC-biased Genotype #1, which was the minor species in the
568 immature *C. sinensis* stroma; the densities of the digested fragments were increased in the maturing
569 and mature stromata, as shown in Figure 13 of (Zhu et al. 2010). Further quantitative analysis
570 demonstrated different maturational alterations of the co-occurring Genotypes #1 and #2 in an
571 asynchronous and disproportional manner in the *C. sinensis* stromata during maturation (Figure 3,
572 modified from Figure 6 of Zhu et al. 2012). The dynamic but asynchronous alteration of the
573 abundance of Genotypes #1 and #2 during *C. sinensis* maturation indicates genomic independence of
574 these 2 *O. sinensis* genotypes but does not support the idea that "Different types of ITS sequences

575 were thus proved to coexist in a single genome" (Li et al. 2013).



576

577 ***Figure 3. Dynamic alterations in the abundance of amplicons of Genotypes #1 and #2 of O. sinensis in the***
578 ***stromata of C. sinensis during maturation*** *(modified from Figure 6 of Zhu* et al. *2012).*
579 *Immature C. sinensis specimens had very short stromata of 1.5 cm on average. Maturing C. sinensis specimens*
580 *had stromata of 4.0 cm on average, without the formation of an expanded fertile portion close to the stromal*
581 *tip. Mature C. sinensis specimens had long stromata of 7.0 cm on average and showed the formation of an*
582 *expanded stromal fertile portion (SFP in red), which was densely covered with numerous ascocarps.*

583 To examine dynamic alterations of the natural biomasses of GC- and AT-biased genotypes of *O.*
584 *sinensis* without PCR manipulation, which may cause nonlinear, exponential amplification of DNA
585 sequences, the Southern blotting technique was applied to examine genomic DNA after enzymolysis
586 with endonuclease *Eco*RI (Zhu et al. 2010). This non-PCR approach identified a single *Eco*RI-
587 sensitive GC-biased DNA moiety in genomic DNA isolated from a pure *H. sinensis* strain (a gift from
588 a senior mycological taxonomist, Prof. Ying-Lan Guo) (*cf.* Figure 4B of Zhu et al. 2010) and a doublet
589 (both *Eco*RI-sensitive Genotype #1 and *Eco*RI-resistant AT-biased genotypes) in genomic DNA
590 isolated from natural *C. sinensis* (Figure 4; reuse of Figure 6 of Zhu et al. 2010 with permission). The
591 lack of *Eco*RI-resistant AT-biased genotypes in pure *H. sinensis* shown in Figure 4B of (Zhu et al.
592 2010) indicates that the sequences of AT-biased genotypes of *O. sinensis* are absent in the genome of
593 GC-biased Genotype #1 *H. sinensis*. The upper panel of Figure 4 shows that Southern blotting without

PCR amplification demonstrated the following:

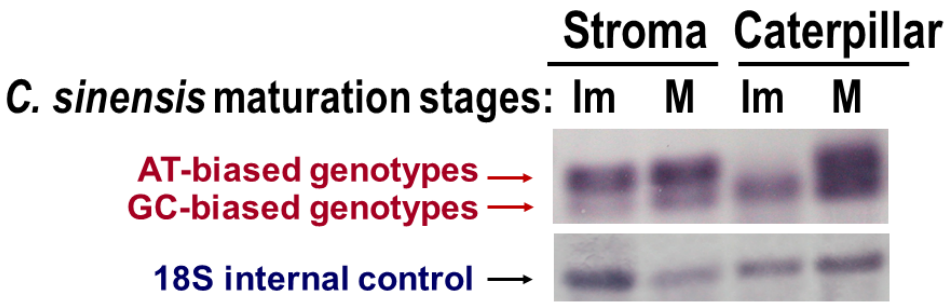


Figure 4. Southern blot of C. sinensis nrDNA in the stroma and caterpillar body of natural C. sinensis during maturation. *[Reproduced with permission from AJBMS (www.nwpii.com/ajbms) Am J Biomed Sci 2010; 2(3): 217-238] [Zhu* et al. *2010].*
Caterpillar refers to caterpillar body. Genomic DNA templates were isolated from the stromata or caterpillar bodies of immature (Im) or mature (M) C. sinensis specimens collected from Sichuan Province in China and prepared using the restriction enzymes EcoRI, DraI and AvaI. ***The upper panel*** *was probed with an H. sinensis-specific probe.* ***The lower panel*** *was probed with a nonspecific 18S internal control probe.*

(1) Comparison of *O. sinensis* genotypes in the stromata: the left 2 lanes of the upper panel of Figure 4 show that the slower-migrating moieties (*Eco*RI-resistant AT-biased genotypes) predominate naturally without PCR amplification in the stromata of immature and mature *C. sinensis*. The faster-migrating moieties (*Eco*RI-sensitive GC-biased *H. sinensis*) are always minor moieties in the *C. sinensis* stromata.

(2) Comparison of *O. sinensis* genotypes in the caterpillar bodies: although the second lane from the right in the upper panel of Figure 4 seems to show a single *Eco*RI-sensitive moiety in the immature caterpillar body but a doublet in the mature caterpillar body shown in the far right lane, PCR experiments identified *Eco*RI-resistant AT-biased genotypes in both samples. These results suggest that the natural quantity of *Eco*RI-resistant AT-biased genotypes in the immature caterpillar body was below the detection limit of Southern blotting, which is much less sensitive than PCR-based techniques. The quantity of AT-biased genotypes was dramatically and disproportionally increased from a micro-quantity in the immature caterpillar body to a level comparable to that of *Eco*RI-sensitive GC-biased *H. sinensis* in the mature caterpillar body.

(3) Comparison of *O. sinensis* genotypes in immature *C. sinensis*: the first and third lanes from the left in Figure 4 show that the abundances and density ratios of *Eco*RI-resistant AT-biased genotypes and *Eco*RI-sensitive GC-biased *H. sinensis* differed significantly in the stroma and caterpillar body of immature *C. sinensis*, which were sampled from the same specimens.

621 (4) Comparison of *O. sinensis* genotypes in mature *C. sinensis*: the first and third lanes from the
622 right in Figure 4 show that the abundances and density ratio of *Eco*RI-resistant and *Eco*RI-
623 sensitive genotypes differed in the stroma and caterpillar body of mature *C. sinensis*, which were
624 sampled from the same specimens.

625 The disproportional distribution of GC- and AT-biased *O. sinensis* genotypes in the stromata and
626 caterpillar bodies of natural *C. sinensis* and their asynchronous alterations of the natural biomasses
627 during *C. sinensis* maturation indicate that the sequences of AT-biased *O. sinensis* genotypes do not
628 reside in the genome of Genotype #1 *H. sinensis*. Instead, the genomically independent AT-biased *O.*
629 *sinensis* genotypes are interindividually accompanied by GC-biased Genotype #1 *H. sinensis* in the
630 stroma and caterpillar body of natural *C. sinensis*.

631 If "Different types of ITS sequences were thus proved to coexist in a single genome" as assumed
632 by Li et al. (2013), the distribution of GC- and AT-biased *O. sinensis* genotypes should be proportional
633 in the *C. sinensis* stroma and caterpillar body and any biomass alterations of the co-occurring
634 genotypes should be synchronous during *C. sinensis* maturation. Southern blotting, as well as RFLP
635 after pretreatment with *Eco*RI enzymolysis and MassARRAY SNP mass spectrometry genotyping
636 assays, clearly demonstrated asynchronous and disproportional alterations of the GC- and AT-biased
637 *O. sinensis* genotypes (Zhu et al. 2010, 2012; Gao et al. 2011, 2012; Yao et al. 2011) and confirmed
638 the genomic independence of the *O. sinensis* genotypes, which belong to independent *O. sinensis*
639 fungi (Stensrud et al. 2007; Xiao et al. 2009).

640 Zhu et al. (2010) and Gao et al. (2011, 2012) demonstrated that Genotype #4 of AT-biased
641 Cluster-B of *O. sinensis* predominated in the immature stroma of natural *C. sinensis,* which was in
642 its asexual growth stage; however, its abundance dramatically declined in the mature stroma, which
643 was in its sexual reproductive stage. The maturational decline of Genotype #4 in the *C. sinensis*
644 stroma was accompanied by a significant reciprocal elevation in the abundance of Genotype #5 of
645 AT-biased Cluster-A of *O. sinensis*, indicating that the multiple AT-biased genotypes belong to
646 independent *O. sinensis* fungi.

647 Li et al. (2022c, 2022d) designed several pairs of genotype-specific primers combined with
648 cloning-based amplicon sequencing and identified Genotypes #5–6 and #16 of AT-biased Cluster-A
649 in the ascospores of natural *C. sinensis*, co-occurring with GC-biased Genotype #1. They identified
650 the sequences of Genotypes #4 and #15 of AT-biased Cluster-B in the stromal fertile portion densely
651 covered with numerous ascocarps but not in ascospores obtained from the same *C. sinensis* specimens,

652 consistent with the findings reported by Li et al. (2013) that AT-biased Genotype #4 was not present
653 in 15 cultures of *C. sinensis* mono-ascospores. Li et al. (2022d) also identified Genotype #13 in
654 semiejected ascospores and Genotype #14 in fully ejected ascospores collected from the same pieces
655 of natural *C. sinensis*. Li et al. (2022c) revealed a high abundance of Genotypes #6 and #15 in the
656 preejection stromal fertile portion of natural *C. sinensis,* but the abundance dramatically declined in
657 the stromal fertile portion after ascospore ejection, while Genotype #5 continued to show high
658 abundance. All the scientific evidence supports the genomic independence hypothesis for multiple *O.*
659 *sinensis* genotypic fungi initially proposed by Xiao et al. (2009).

660 After the hypothesis of "ITS pseudogenes … in a single genome" was published and subjected
661 to academic discussion, Li et al. (2013) had sufficient time to validate their hypothesis by conducting
662 genome sequencing on one of the Group 3 strains (*cf.* Section III-3); however, scientific papers
663 addressing this have yet to appear. Instead, Li et al. (2013, 2017, 2019, 2020c) stated repeatedly that
664 "Different types of ITS sequences were thus proved to coexist in a single genome", "from a single
665 genome, and not from different genomes", and "The coexistence of diverged ITS paralogs in a single
666 genome". These scientists either did not conduct genome sequencing of one of the Group 3 strains or
667 may have failed the experiments due to difficulties in assembly of shotgun reads to experimentally
668 verify their hypothesis that Genotypes #1 and #5 coexist in "a single genome", although this genome
669 sequencing strategy may not be able to confirm the coexistence of AT-biased Genotypes #4, #6 and
670 #15–17 as the genomic components of "a single genome" of GC-biased Genotype #1 *H. sinensis*.

671 Bushley et al. (2013) and Hu et al. (2013) hypothesized self-fertilization as the reproductive mode
672 of Genotype #1 *H. sinensis* under homothallism or pseudohomothallism based on the identification
673 of mating-type genes of both *MAT1-1* and *MAT1-2* idiomorphs in 2 *H. sinensis* strains. Li et al. (2020b)
674 stated accordingly that "*O. sinensis* is one of the only known homothallic species in
675 *Ophiocordycipitaceae* … the homothallic life-style may have increased the likelihood and frequency
676 of sex in *O. sinensis*, leading to increased RIP in this species." However, Zhang and Zhang (2015)
677 revealed differential occurrence of the mating-type genes of *MAT1-1* and *MAT1-2* idiomorphs in
678 numerous *H. sinensis* strains and alternatively hypothesized facultative hybridization as the mode of
679 *O. sinensis* reproduction.

680 Li et al. (2022b, 2022e) reported significant intraspecific genetic variations among *H. sinensis*
681 strains and differential expression of *H. sinensis* genes. Li et al. (2021) further reviewed the genetic
682 information of 236 *H. sinensis* strains and confirmed the differential occurrence of mating-type genes

683 of the *MAT1-1* and *MAT1-2* idiomorphs: 22 strains contained only MAT1-1-1 genes, 65 possessed
684 only MAT1-2-1 genes, and 149 had both genes, consistent with the findings reported by Zhang and
685 Zhang (2015). Li et al. (2021) identified differential but reciprocal transcription of mating-type genes
686 in two *H. sinensis* strains. The MAT1-1-1 transcript was absent, but the MAT1-2-1 transcript was
687 present in transcriptome assembly GCQL00000000 of *H. sinensis* Strain L0106 (Liu et al. 2015; Li
688 et al. 2021). On the other hand, the MAT1-1-1 gene was shown to be transcribed with normal splicing
689 of both intron I and II in *H. sinensis* Strain 1229, but the MAT1-2-1 transcript included an unspliced
690 intron I, which contains 3 stop codons (Bushley et al. 2013; Li et al. 2021). Bushley et al. (2013)
691 stated that the total "RNA was treated with DNase I … until no DNA contamination could be detected"
692 prior to reverse transcription and cDNA library construction, and "the products (*i.e.*, cDNA) were
693 derived from mRNA and did not result from contamination of genomic DNA. In contrast, comparison
694 of DNA and cDNA sequences of MAT1-2-1 revealed only one spliced intron of 55 bp (Fig S2)." Thus,
695 the special phenomenon of MAT1-2-1 transcription reported by Bushley et al. (2013) indicates the
696 production of a largely truncated and dysfunctional MAT1-2-1 protein lacking the majority of the
697 protein, encoded by exons II and III. Thus, the transcriptional silencing of one of the mating-type
698 genes and the MAT1-2-1 transcript with an unspliced intron I constitute *O. sinensis* reproduction
699 controls at the transcriptional level and the coupled transcriptional-translational level.

700 Accordingly, Li et al. (2021) concluded self-sterility of GC-biased Genotype #1 *H. sinensis*,
701 suggesting a requirement of sexual partners for reproduction. These partners could be either
702 intraspecific species, such as *H. sinensis* Strains L0106 and 1229, for physiological heterothallism
703 regardless of whether *H. sinensis* is monoecious or dioecious, or heterospecific fungal species that
704 may undergo hybridization if they are able to overcome interspecific reproductive barriers (Zhang
705 and Zhang 2015; Li et al. 2021, 2022d, 2022e). Because the taxonomic positions of Genotypes #2–
706 17 of *O. sinensis* have not been determined, it is debatable whether the coexisting genotypes represent
707 intraspecific species involved in physiological heterothallism or heterospecific species involved in
708 fungal hybridization. Thus, the notion of "the homothallic life-style in *O. sinensis* … leading to
709 increased RIP", put forth by Li et al. (2020b), may be groundless because "the homothallic life-style
710 in *O. sinensis*" is scientifically incorrect and "increased RIP" that theoretically induces C-to-T and
711 G-to-A transition mutations was not confirmed based on the analysis of the multiple repetitive ITS
712 copies in the genomes of Genotype #1 *H. sinensis* strains, as analyzed in Section III-1.
713

III-5. Transcriptional silencing of 5.8S genes

In addition to the identification of both Genotypes #1 and #5 ITS sequences from genomic DNA isolated from 8 of 15 cultures of multicellular heterokaryotic mono-ascospores, Li et al. (2013) examined the 5.8S gene transcripts of these clones. They reported the identification of 5.8S cDNA for Genotype #1, but not for Genotype #5, in a cDNA library constructed from total ascosporic RNA extracted from Clone 1220, one of the 8 genetically heterogeneous clones. However, the authors overinterpreted that the 5.8S genes of all AT-biased genotypes were permanently dysfunctional pseudogenic components coexisting in "a single genome" with a functional copy of GC-biased Genotype #1, although they did not detect the ITS sequences of AT-biased Genotypes #4, #6 and #15–17 in the genomic DNA isolated from any of the ascosporic clones. The logical reasoning underlying the dysfunctional pseudogene hypothesis was explained by Li et al. (2019) as follows: "If AT-biased ITS sequences are not pseudogenes, 5.8S cDNA should have been detected because rRNAs are critical and essential for protein synthesis in all living organisms", which touches on 2 scientific issues. First, are the 5.8S genes of all genotypes of *O. sinensis* actively transcribed in *C. sinensis* ascospores (discussed in this section)? Second, is it correct to conclude that "5.8S cDNA should have been detected" by the techniques designed and used by Li et al. (2013) (discussed in Section III-6)?

Xiang et al. (2014) reported a metatranscriptome study of natural *C. sinensis* specimens collected from Kangding County of Sichuan Province, China. The GAGW00000000 metatranscriptome sequences do not contain any 5.8S gene transcripts of *C. sinensis*-associated fungi (including Genotype #1 and all other genotypes of *O. sinensis*), indicating transcriptional silencing of 5.8S genes of all cocolonized fungi in natural settings.

Xia et al. (2017) reported another metatranscriptome study of mature *C. sinensis* specimens collected from Deqin County of Yunnan Province, China. The assembled metatranscriptome sequences were not uploaded to GenBank but instead to another depository server www.plantkingdomgdb.com/Ophiocordyceps_sinensis/data/cds/Ophiocordyceps_sinensis_CDS.fas. Similarly, this assembled metatranscriptome contains no 5.8S gene transcripts. However, the authors uploaded the unassembled shotgun reads (SRR5428527) to GenBank. A BLAST search of the SRR5428527 Shotgun Read Archive (SRA) using the MegaBLAST algorithm identified 1894 metatranscriptome shotgun reads (100 bp each), which were 100% homologous to the 5.8S gene sequences (156 bp) of GC-biased Genotype #1 *H. sinensis*.

It is well established that fungal genomes contain multiple copies of rDNA ITS1-5.8S-ITS2

745 sequences (Tables 1 and 3; Figures 1–2). The 1894 unassembled metatranscriptome shotgun reads
746 (SRR5428527) of natural *C. sinensis* may be derived from highly homologous nonrDNA sequences
747 of slowly evolved genes within fungal metagenomes, while there are no rDNA 18S and 28S gene
748 segments of *H. sinensis* homologous to the 5.8S gene. Figure 5 shows the alignment of the 5.8S gene
749 (the query sequence; segment 218–373 of Genotype #1 *H. sinensis* AB067721) and overlap of the
750 first 75 of the 1894 metatranscriptome shotgun reads of mature natural *C. sinensis*. The short length
751 (100 bp) of SRA shotgun reads resulted in overlap between the 1894 fragmented sequence reads and
752 transcripts of nonrDNA genes (further discussion below).

753

754 ***Figure 5. Alignment of the query sequence in green (the 5.8S gene of Genotype #1 H. sinensis AB067721)***
755 ***and the 75 of 1894 overlapping (pink bars) shotgun reads (GenBank SRR5428527) of mature C. sinensis.***

756 Various degrees of similarity were found when the 1894 unassembled metatranscriptome shotgun
757 reads were aligned with the 5.8S gene sequences of the 17 genotypes of *O. sinensis*. Table 4 shows 6
758 examples of shotgun reads of SRX2718557 aligned with the 5.8S sequences of *O. sinensis* genotypes,
759 *Paecilomyces hepiali*, an AB067719-type Group-E fungus, and *Tolypocladium sinensis* (Group D)
760 (Stensrud et al. 2007). Obviously, the high homology between the shotgun reads and the various 5.8S
761 gene sequences may not conclusively determine their fungal sources in natural *O. sinensis*, which
762 contains >90 species spanning >37 fungal genera (Zhang et al. 2010, 2018; Xiang et al. 2014; Xia et
763 al. 2015, 2017). Because no 5.8S transcript was identified from the assembled metatranscriptome
764 sequences in the same study, the shotgun reads might not be derived from the 5.8S genes of multiple
765 cocolonizing fungi, including the multiple genotypes of *O. sinensis*. Liu et al. (2015) reported the
766 mRNA transcriptome assembly GCQL00000000 derived from *H. sinensis* Strain L0106, which
767 contained no transcript of rRNA or tRNA. Continuous *in vitro* liquid fermentation performed in this
768 study induced a nonlinear reduction in the total number of transcriptomic unigenes, first from 25,511
769 (after 3 days of fermentation) to 25,214 (after 6 days of fermentation) and then more sharply to 16,245
770 (after 9 days of fermentation), indicating the nonnatural switching on or off of many genes. The
771 average length of transcripts of Strain L0106 increased nonlinearly, with a gradual decrease in the
772 GC content from 681 nt (GC=58.2%) (3 days) to 682 nt (GC=57.9%) (6 days) and then a dramatic
773 increase to 994 nt (GC=57.0%) (9 days). The differential transcriptional activation and silencing of
774 *H. sinensis* genes were caused by continuous *in vitro* liquid fermentation of *H. sinensis* Strain L0106.
775 Presumably, the prolonged 25-day liquid incubation period adopted by Li et al. (2013) might have a
776 greater impact on the unbalanced transcription of long and short fungal genes with various GC content
777 biases.

778 Apparently, Li et al. (2013, 2017, 2019, 2020b) neglected the evidence of transcriptional silencing
779 of 5.8S genes demonstrated in metatranscriptome studies in natural settings and that of dynamic and
780 unbalanced transcriptional activation and silencing of numerous functional genes during the
781 continuous *in vitro* fermentation or liquid incubation of Genotype #1 *H. sinensis*. Because of the
782 conservative evolution of 5.8S genes and many other nonrDNA genes and the incomplete study
783 design and controversial findings reported by Li et al. (2013), it needs to be verified (probably using
784 other sophisticated transcriptional expression methods) whether the identified cDNA was truly
785 derived from the 5.8S gene of Genotype #1 *H. sinensis*, or from mutant genotypes of *O. sinensis* or

786 other colonized fungi, or even from nonrRNA genes (further discussion below) before confidently
787 and reasonably declaring that the 5.8S genes of the multiple AT-biased genotypes of *O. sinensis* are
788 truly permanently nonfunctional pseudogenes under the "ITS pseudogene … in a single genome"
789 hypothesis.

790 ***Table 4.*** *Percent similarity between the unassembled metatranscriptome shotgun reads (SRA:*
791 *SRR5428527) of natural C. sinensis and the 5.8S genes of multiple genotypes of O. sinensis and*
792 *Paecilomyces hepiali, an AB067719-type fungus (Group-E) and Tolypocladium sinensis (Group D).*

Geno-type	GenBank Accession #	% similarity between 5.8S genes and *C. sinensis* shotgun sequences					
		SRR5428527.17546327.1	SRR5428527.17554054.1	SRR5428527.17565575.2	SRR5428527.17679995.2	SRR5428527.17911498.2	SRR5428527.17929569.2
#1	AB067721	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
#2	MG770309	96%	96%	**100%**	96%	**98%**	96%
#3	HM595984	**100%**	**99%**	**99%**	**100%**	**99%**	**100%**
#7	AJ488254	**98%**	**98%**	**99%**	**98%**	**98%**	**98%**
#8	GU246286	94%	96%	95%	94%	**97%**	94%
#9	GU246288	**99%**	**99%**	**99%**	**100%**	**99%**	**100%**
#10	GU246287	96%	94%	93%	95%	93%	**97%**
#11	JQ695935	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
#12	GU246296	**99%**	**99%**	**99%**	**99%**	**99%**	**99%**
#13	KT339190	**98%**	93%	93%	**98%**	93%	**98%**
#14	KT339178	**98%**	93%	93%	**98%**	93%	**98%**
#4	AB067744	86%	87%	87%	83%	87%	83%
#5	AB067740	85%	87%	87%	84%	87%	84%
#6	KJ720572	85%	85%	87%	84%	86%	84%
#15	KT232017	85%	88%	88%	84%	88%	84%
#16	KT232019	86%	86%	87%	85%	87%	85%
#17	KT232010	87%	88%	88%	87%	88%	87%
P.h.	EF555097	**99%**	**98%**	**97%**	**99%**	**98%**	**99%**
	AB067719	**98%**	93%	93%	**98%**	93%	**98%**
T.s.	AB067720	**98%**	**97%**	**96%**	**98%**	**97%**	**98%**

793 *Note: GC-biased genotypes are listed in blue, and AT-biased genotypes are listed in red. GenBank*
794 *sequence ID SRA: SRR5428527 fragments (17546327.1, 17554054.1, 17565575.2, 17679995.2,*
795 *17911498.2, and 17929569.2) are the unassembled transcriptome shotgun sequences obtained from*
796 *the mature C. sinensis specimen collected from Deqin County of Yunnan Province by Xia* et al. *(2017).*
797 *P.h. refers to Paecilomyces hepiali; AB067719 refers to an AB067719-type Group-E fungus; T.s.*
798 *refers to Group-D Tolypocladium sinensis (Stensrud* et al. *2007). Highly homologous percent*
799 *similarities (97%–100%) are bolded in* ***green****.*

800

801 ## III-6. Identification and analysis of 5.8S cDNA using 5.8S-F/R primers

802 In addition to the discussion above of the transcriptional silencing of 5.8S genes in natural
803 settings and that of the significant impact of continuous *in vitro* liquid-phase culture on unbalanced
804 gene transcription, the second issue related to the logic given by Li et al. (2019) is whether the
805 methods designed by Li et al. (2013) were scientifically capable of identifying 5.8S gene cDNAs of
806 multiple *O. sinensis* genotypes (Li et al. 2020a).

807 For instance, Li et al. (2019) referred to "The 5.8S-F/R primers … located in the most conserved
808 region", which may reduce the primers' specificity for 5.8S cDNA because of the conservative
809 evolution of the 5.8S gene and many other genes. Figure 6 shows that primer 5.8S-F is 95% (20/21
810 nt) similar and 100% (21/21 nt) identical to the 5.8S transcript sequence of AT-biased Genotypes #4–
811 5, respectively. Primer 5.8S-R is 90% (18/20 nt, with 2 flanking residues downstream in the ITS2
812 region) and 95% (19/20 nt) similar to the 5.8S transcript sequence of AT-biased Genotypes #4–5,
813 respectively. In contrast to the 95% (20/21 nt) and 85% (17/20 nt) similarities of primers 5.8S-F/R to
814 the 5.8S transcript sequence of GC-biased Genotype #1, it is apparent that Li et al. (2013) designed
815 the 5.8S-F/R primers based on the AT-biased 5.8S gene sequences, making it easier to amplify the
816 5.8S cDNA sequence of Genotype #5 than Genotype #1. The primers 5.8S-F/R showed variable
817 similarity, ranging from 61.9%–100% and 31.8%–95.5%, respectively, among all 17 genotypes of *O.*
818 *sinensis*, preferentially amplifying the most easily amplified cDNA sequence and consequently
819 causing false-negative results for other 5.8S cDNA sequences.

820 Section III-5 and Table 4 show 100% homology between the 1894 unassembled
821 metatranscriptome shotgun reads (100 bp) and the 5.8S gene sequence (156 bp) of GC-biased
822 Genotype #1 *H. sinensis*, probably including a large number of nonrRNA transcripts among the 1894
823 shotgun reads. A BLAST search of the GenBank database showed that the 5.8S-F/R primers were
824 100% homologous to mRNA transcriptome assemblies GCQL01001191, GCQL01006777,
825 GCQL01007818, and GCQL01013443 of Genotype #1 Strain L0106, which contains no rRNA
826 transcripts and assembled metatranscriptome assemblies GAGW01003691, GAGW01007670, and
827 GAGW01007671 of natural *C. sinensis* (Li et al. 2020a). Alignment analysis showed that
828 metatranscriptome GAGW01003691 (2984 bp) was 99.8% homologous to a segment of mRNA
829 transcriptome GCQL01006777 of *H. sinensis* Strain L0106. Metatranscriptomes GAGW01007670
830 (234 bp) and GAGW01007671 (1620 bp) are 100% identical to segments of GCQL01007818 and

GCQL01001191, respectively. Because mRNA transcriptome assembly GCQL00000000 contains no rRNA transcripts, the extremely high homology between the metatranscriptome and transcriptome sequences indicates that the 5.8S-F/R primers designed by Li et al. (2013) may nonspecifically amplify nonrRNA components in the cDNA library derived from Group 3 clones (*cf*. Section III-3).

```
GT1 AB067721 1     TAGAGGAAGTAAAAGTCGTAACAAGGTCTCCGTTGGTGAACCAGCGGAGGGATCATTATC  60
GT4 AB067744 1         -------------------------------A--A---T-----------T------  56
GT5 AB067740 1     --A------------------T------------A--A---T-----------T------  60

GT1 AB067721 61    GAGTCACCACTCCCAAACCCCCTGCGAACACCACAGCAGTTGCCTCGGCGGGACCGCCCC  120
GT4 AB067744 57    ------------C--------------A--C------------------------------  116
GT5 AB067740 61    ----T--T----TT---------ATA--T--T-T-----------TA------------T  120

GT1 AB067721 121   GGCGCCCCAGGGCCCGGACCAGGGCGCCCGCCGGAGGACCCCCAGACCCTCCTGTCGCAG  180
GT4 AB067744 117   -------T-----------T---------------------TT----------A---T--  176
GT5 AB067740 121   A------T-A---------T-AA------------------TT----------A-TAT--  180

GT1 AB067721 181   TGGCATCTCTCAGTCAAGAAGCAAGCAAATGAATCAAAACTTTCAACAACGGATCTCTTG  240
GT4 AB067744 177   -A--------T---T---------------A---T-A-ACTTTTAATAACGGATCTCTTA  236
GT5 AB067740 181   -A-T------T---T---------------A---T---ACTTTTAACAACGGATCTCTTA  240
Primer 5.8S-F                                            ACTTTTAACAACGGATCTCTT

GT1 AB067721 241   GTTCTGGCATCGATGAAGAACGCAGCGAAATGCGATAAGTAATGTGAATTGCAGAATTCA  300
GT4 AB067744 237   ---T-A----TA--A----------------A-------------A-----T------T-  296
GT5 AB067740 241   -----A----TA--A--A---------------A---------A-A----A-------T-  300

GT1 AB067721 301   GTGAACCATCGAATCTTTGAACGCACATTGCGCCCGCCAGCACTCTGGCGGGCATGCCTG  360
GT4 AB067744 297   --A---T-----------A----T-T-C-A-------T--------A------ATGCCTG  356
GT5 AB067740 301   --A---T-----------A------T---A-------T--------A------ATGCCTA  360
Primer 5.8S-R                                                        ATGCTTA

GT1 AB067721 361   TCCGAGCGTCATCTCAACCCTCGAGCCCCCCGCCTCGCGGCGGCGGGGCCCGGCCTTGGG  420
GT4 AB067744 357   TCCGAGCGTTACCTT------TA----------------------------------A--  416
GT5 AB067740 361   TCCGAGCGTTATCTT----------------------------------T-------A--  420
Primer 5.8S-R      TCCGAGCGTTATCTT

GT1 AB067721 421   GGTCACGGCCCCGCGCCGCCCCCTAAACGCAGTGGCGACCCCGCCGCGGCTCCCCTGCGC  480
GT4 AB067744 417   ------------A--------------------A-T--------------------A---  476
GT5 AB067740 421   ---T------------------T------T---A----------------------A--T  480

GT1 AB067721 481   AGTAGCTCGCTGAGAACCTCGCACCGGGAGCGCGGAGGCGGTCACGCCGTGAAACCACCA  540
GT4 AB067744 477   -A-----TA--A------------TA------------------------A-------T-  536
GT5 AB067740 481   G-------A-C----------------A----------------------A-------T-  540

GT1 AB067721 541   CACCCTCCAGTTGACCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCAATAA  600
GT4 AB067744 537   T------TG---A---------T-----------------A-------------------  596
GT5 AB067740 541   -G--T-TT----A---------T-----------------A-------------T-----  600

GT1 AB067721 601   GCGGAGGAAAAGAAACCAACAGGGATTGCCCCAGTAACGGCGAGTGAAGCGGCAACAGCT  660
GT4 AB067744 597   --------                                                       604
GT5 AB067740 601   ----------------T--T----------TT-------------A------T--T----  660
```

Figure 6. ITS sequence alignment of AB067721 of GC-biased Genotype #1, AB067744 and AB067740 of AT-biased Genotypes #4–5 of O. sinensis and the 5.8S-F/R primers.
"GT1", "GT4" and "GT5" represent Genotypes #1 and #4–5 of O. sinensis, respectively. The sequences in blue are 18S and 28S genes (5' and 3' end regions, respectively). Those in pink represent sequences of 5.8S genes. Those in black are ITS1 between the 18S and 5.8S genes and ITS2 between the 5.8S and 28S genes. The sequences underlined in green are for primers 5.8S-F/R, while primer 5.8S-R was converted to the sense chain sequence. Hyphens indicate identical bases.

Li et al. (2013) did not disclose the sequences of 5.8S cDNA that they identified. Instead, they

844 reported that "Of these clones, 28 showed identical sequences (104 bp excluding the primers) to the
845 most frequent sequence in group A (JQ900148). The remaining 14 sequences … had single
846 substitutions at various positions", where group A = Genotype #1 *H. sinensis*. GenBank annotates the
847 location of the 5.8S gene of JQ900148 as nucleotides 192→346, covering 155 nt (not the multiple of
848 three), which is not in accord with the open reading frame of the gene.

849 Alternatively, the 5.8S genes are annotated with nucleotide positions of 218→373 in AB067721
850 of Genotype #1 and AB067740 of Genotype #5, and 214→369 in AB067744 of Genotype #4,
851 covering 156 nt and 52 codons in the 3 cases (Zhu & Li 2017; Li et al. 2020a, 2022e). Excluding the
852 5.8S-F/R primer regions, the remaining sequence of the 5.8S gene consists of 114 bp, with 10
853 additional nucleotides (mismatches) beyond the 104 bp of all cDNA clones reported by Li et al. (2013).
854 Adding "single substitutions at various positions" in 14 other clones reported by Li et al. (2013), the
855 42 clones presumably contained either 10 or 11 mismatches (8.8–9.6%) in the 114-bp segment of the
856 5.8S gene cDNA, corresponding to 90.4%–91.2% sequence similarity to the 5.8S gene sequence of
857 Genotype #1 of *O. sinensis*. Thus, attributing such high dissimilarity rates to "mismatches during
858 reverse transcription or PCR amplification" may not provide a convincing explanation, giving rise to
859 the question of whether the identified "5.8S cDNA" with a large number of mismatches was truly
860 derived from the 5.8S gene of *O. sinensis* Genotype #1, as assumed by Li et al. (2013, 2019, 2020b).
861 Considering the intensive epigenetic methylation of cytosine residues in a 5S RNA gene that is
862 associated with the transcriptional silencing of GC-biased genes, as reported by Selker and his
863 colleague (1985, 1990, 1997, 2002) and Gladyshev (2017), it may become legitimate to doubt
864 whether the 5.8S cDNA sequences that were identified by Li et al. (2013) truly belong to GC-biased
865 Genotype #1; this genotype contains more cytosine residues and is more susceptible to epigenetic
866 methylation attack and subsequent translational silencing than AT-biased Genotype #5. Accordingly,
867 it is reasonable to doubt the "ITS pseudogene" hypothesis, which postulates that the 5.8S genes of
868 AT-biased genotypes are permanently dysfunctional, rather than GC-biased Genotype #1.

869 Prior to declaring that the 5.8S genes of all AT-biased *O. sinensis* genotypes are nonfunctional
870 pseudogenes co-occurring with a functional 5.8S gene in the genome of GC-biased Genotype #1,
871 several other concerns should be addressed.

872 (1) Li et al. (2013) did not detect the ITS sequences of Genotypes #2–4 and #6–17 of *O. sinensis* in
873 the genomic DNA pool of ascosporic cultures, indicating that inferring that the 5.8S genes of
874 these genotypes are dysfunctional "ITS pseudogene" components of the genome of Genotype #1

875 *H. sinensis* is an apparent overinterpretation.

876 (2) The sequences of Genotypes #2–17 are absent in the genome assemblies ANOV00000000,
877 JAAVMX000000000, LKHE00000000, LWBQ00000000 and NGJJ00000000 of *H. sinensis*
878 Strains Co18, IOZ07, 1229, ZJB12195 and CC1406-2031229, respectively (Hu et al. 2013; Li et
879 al. 2016a; Jin et al. 2020; Liu et al. 2020; Shu et al. 2020), indicating genomic independence of
880 the genotypes of *O. sinensis*.

881 (3) Selker and his colleague (1985, 1990, 1997, 2002) and Gladyshev (2017) have repeatedly
882 reported epigenetic methylation of nearly all cytosine residues in 5.8S RNA genes and
883 subsequent transcription silencing of GC-biased genes, while the 5.8S gene of GC-biased
884 Genotype #1 contains more cytosine residues and is more susceptible to epigenetic methylation
885 attack and subsequent translational silencing than 5.8S genes of AT-biased genotypes.

886 (4) The culture-dependent mycological approach used by Li et al. (2013) might have missed some
887 fungal species that are nonculturable or difficult to culture under the *in vitro* experimental settings
888 of the study. The culture-dependent strategy might have a significant impact on the transcription
889 of many fungal genes, with many genes being switched on or off nonnaturally, as evidenced by
890 the nonlinear reduction in the total number of transcriptomic unigenes and increases in the
891 average length but decreases in the GC content of transcripts during 3–9 days of liquid
892 fermentation of *H. sinensis* Strain L0106 (Liu et al. 2015). A much greater impact on differential
893 transcriptional activation and silencing of many genes may be expected after the prolonged 25-
894 day liquid incubation period adopted by Li et al. (2013).

895 (5) Three distinct types of secondary steric conformations of 5.8S rRNA were predicted for Groups
896 A–C (*i.e.*, Genotypes #1 and #4–5) of *O. sinensis* by Li et al. (2013). The possibility of producing
897 circular RNA through back splicing-exon circularization (Dong et al. 2014a) should be
898 considered during study design because these distinct steric structures may have a considerable
899 impact on reverse transcription PCR and cDNA sequencing. In addition, the ITS1-5.8S-ITS2
900 sequences of Genotypes #2 and #6 may adopt different types of steric conformations (Zhu & Wu
901 2015; Li et al. 2016b, 2022e; Zhu & Li 2017); thus, the design of other genotype-specific primers
902 and the combined use of other molecular techniques should be considered.

903 (6) Wei et al. (2016) reported the identification of a single teleomorph of AT-biased Genotype #4 of
904 *O. sinensis* in cultivated *C. sinensis* and a single teleomorph of GC-biased Genotype #1 in natural

905 *C. sinensis* (*cf*. Figure 6 of Wei et al. 2016). If AT-biased Genotype #4 represents a nonfunctional
906 ITS pseudogene, as believed by Li et al. (2013, 2017, 2019, 2020b), the single *O. sinensis*
907 teleomorph in cultivated *C. sinensis* might have disturbed teleomorphic functions, leading to
908 reproductive sterility and an abnormal lifecycle of cultivated *C. sinensis*.

909 (7) Li et al. (2022c, 2022e) reported the identification of teleomorphic Genotypes #4 and #15 of *O.*
910 *sinensis* of AT-biased Cluster-B in the stromal fertile portion densely covered with numerous
911 ascocarps but not in ascospores collected from the same pieces of natural *C. sinensis*. AT-biased
912 Genotypes #6 and #15 were found at high abundance in the stromal fertile portion prior to
913 ascospore ejection, and their abundance drastically declined after ascospore ejection, while
914 teleomorphic Genotype #5 maintained its high abundance in the stromal fertile portion before
915 and after ascospore ejection (Li et al. 2022a, 2022e).

916 (8) The 5.8S genes of multiple *O. sinensis* genotypes may be transcriptionally activated or silenced
917 in different developmental and maturational stages of natural *C. sinensis*. Significant changes in
918 the proteomic expression profiles of the stroma and caterpillar body of natural *C. sinensis* during
919 maturation constitute evidence of such dynamic alterations in epigenetic, transcriptional,
920 translational and posttranslational modifications of numerous genes (Dong et al. 2014b).

921 Apparently, the techniques used by Li et al. (2013) provided insufficient and nonconclusive
922 evidence to support their "ITS pseudogenes" hypothesis. The whole package of gene expression in
923 *O. sinensis* fungi and natural *C. sinensis* needs to be further explored before the 5.8S genes of AT-
924 biased genotypes of *O. sinensis* can be defined as permanently nonfunctional pseudogenes.

925

926 VI. Conclusion

927 This paper analyzes the data related to the multiple repetitive ITS copies in the genome of
928 Genotype #1 *H. sinensis* and multiple genotypes of *O. sinensis* fungi and discusses the "ITS
929 pseudogene … in a single genome" hypothesis for all AT-biased *O. sinensis* genotypes (Li et al. 2013,
930 2017, 2019, 2020b). The repetitive ITS copies in the genome of Genotype #1 contain multiple
931 scattered insertion, deletion and transversion point mutations and may not be generated through RIP
932 mutagenesis because RIP theoretically causes C-to-T and G-to-A transitions. The GC-biased
933 repetitive ITS copies are genetically and phylogenetically distinct from AT-biased *O. sinensis*
934 genotypes that contain multiple scattered transition alleles. The AT-biased genotypes are genomically

independent from the genome of Genotype #1 and might have been generated from a common ancestor through RIP or other mutagenic and evolutionary mechanisms in the long history of evolution and become independent *O. sinensis* fungi with independent genomes coexisting with GC-biased genotypes in natural *C. sinensis*. Li et al. (2013) provided incomplete and controversial information to support their "ITS pseudogene … in a single genome" hypothesis, which needs to be further validated.

V. Acknowledgment:

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.

References

[1] Bushley, K.E., Li, Y., Wang, W.-J., Wan,g X.-L., Jiao, L., Spatafora, J.W., Yao, Y.-J., 2013. Isolation of the MAT1-1 mating type idiomorph and evidence for selfing in the Chinese medicinal fungus *Ophiocordyceps sinensis*. Fungal Biol. 117, 599–610. (PMID: 22181845 DOI: 10.1016/j.funbio.2013.06.001)

[2] Chen C-S, Hseu R-S, Huang C-T. 2011. 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. (www.intechopen.com)

[3] Chen, Y.-Q., Hu, B., Xu, F., Zhang, W., Zhou, H., Qu, L.-H., 2004. Genetic variation of *Cordyceps sinensis*, a fruit-body-producing entomopathogenic species from different geographical regions in China. FEMS Microbiol. Lett. 230, 153–158. (PMID: 14734179 DOI: 10.1016/S0378-1097(03)00889-9)

[4] Dong, R., Che,n L.-L., Yang, L., 2014a. Circular RNA studies in the post-genomic era. Chin. J. Cell Biol. 36 (11), 1455–1459. (DOI: 10.11844/cjcb.2014.11.9001)

[5] 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., 2014b. Altered proteomic polymorphism in the caterpillar body and stroma of natural *Cordyceps sinensis* during maturation. PLoS ONE. 9 (10), e109083. (PMID: 25310818 doi:10.1371/journal.pone.0109083)

[6] Engh, I.B., 1999. Molecular phylogeny of the *Cordyceps-Tolypocladium* complex. Candidate scientific thesis, Department of Biology, University of Oslo, Oslo. 1999.

[7] Gao, L., Li, X.-H., Zhao, J.-Q., Lu, J.-H., Zhu, J.-S., 2011. Detection of multiple *Ophiocordyceps sinensis* mutants in premature stroma of *Cordyceps sinensis* by MassARRAY SNP MALDI-TOF mass spectrum genotyping. Beijing Da Xue Xue Bao. 43 (2), 259–266. (PMID: 21503123 doi: 10.3969/j.issn.1671-167X.2011.02.019)

[8] Gao, L., Li, X.-H., Zhao, J.-Q., Lu, J.-H., Zhao, J.-G., Zhu, J.-S., 2012. 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*. Beijing Da Xue Xue Bao. 44 (3), 454–463. (PMID: 22692321 doi: 10.3969/j.issn.1671-167X.2012.03.025)

[9] Gladyshev E. 2017. Repeat-Induced Point Mutation (RIP) and Other Genome Defense Mechanisms in Fungi. Microbiol. Spectr. 5(4): . (PMID: 28721856 doi: 10.1128/microbiolspec.FUNK-0042-2017)

[10] Hawksworth DL, Crous PW, Redhead SA, Reynolds DR, Samson RA, Seifert KA, Taylor JW, Wingfield MJ, Abaci O, Aime C, Asan A, Bai F-Y, de Beer ZW, Begerow D, Berikten D, Boekhout T, K Buchanan P, Burgess T, Buzina W, Cai L, Cannon PF, Crane JL, Damm U, Daniel H-M, van Diepeningen AD, Druzhinina I, Dyer PS, Eberhardt U, Fell JW, Frisvad JC, Geiser DM, Geml J, Glienke C, Gräfenhan T, Groenewald JZ, Groenewald M, de Gruyter J, Guého-Kellermann E, Guo L-D, Hibbett DS, Hong S-B, de Hoog GS, Houbraken J, Huhndorf SM, Hyde K D, Ismail A, Johnston PR, Kadaifciler DG, Kirk PM, Kõljalg U, Kurtzman CP, Lagneau P-E, Lévesque CA, Liu X, Lombard L, Meyer W, Miller A, Minter DW, Najafzadeh MJ, Norvell L, Ozerskaya SM, Oziç R, Pennycook SR, Peterson SW, Pettersson OV, Quaedvlieg W, Robert VA, Ruibal C, Schnürer J, Schroers H-J, Shivas R, Slippers B, Spierenburg H, Takashima M, Taşkın E, Thines M, Thrane U, Uztan AH,van Raak M, Varga J, Vasco A, Verkley G, Videira SIR, de Vries RP, Weir BS, Yilmaz N, Yurkov A, Zhang N. The Amsterdam declaration on fungal nomenclature. IMA Fungus 2011; 2: 105−112 (PMID: 22679594 DOI: 10.5598/imafungus.2011.02.01.14).

[11] Hu, X., Zhang, Y., Xiao, G., Zheng, P., Xia, Y., Zhang, X., St Leger, R.J., Liu, X., Wang, C., 2013. Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus. Chin. Sci. Bull. 58, 2846–2854. (doi: 10.1007/s11434-013-5929-5)

[12] Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformat. 17: 754−755. (PMID: 11524383 DOI: 10.1093/bioinformatics/17.8.754)

[13] 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. 2020. Genome sequencing and analysis of fungus *Hirsutella sinensis* isolated from *Ophiocordyceps sinensis*. AMB Expr. 10:105. (PMID: 32494871 DOI: 10.1186/S13568-020-01039-x)

[14] Kinjo, N., Zang, M., 2001. Morphological and phylogenetic studies on *Cordyceps sinensis* distributed in southwestern China. Mycoscience. 42, 567–574. (https://doi.org/10.1007/BF02460956)

[15] Li C-L, Tan N-Z, Berger JL, Ferguson SB, Zhang Y, Prolla TA, Zhu J-S. 2011. The Combined use of

1000 ageLOC whole-gene expression profiling technology and mouse lifespan test in anti-aging herbal product
1001 study. Proc. 2011 New TCM Prod. Innovat. Indust. Devel. Summit, Hangzhou, China, 2011, pp. 443–448.
1002 (https://xueshu.baidu.com/usercenter/paper/show?paperid=08341c17fa58c8f85584b92572b90f75&site=xueshu_se)
1003 [16] Li X-Z, Li Y-L, Zhu J-S. *Differential expression of mating-type genes in Hirsutella sinensis and natural*
1004 *Cordyceps sinensis*. Proc. 2021 Intl. Conf. Info. Technol. Biomed. Engineering, 2021, pp. 318−333.
1005 (https://conferences.computer.org/icitbepub/pdfs/ICITBE2021-6JNtJiThlp10s9N5Sodvuy/009900a318/009900a318.pdf)
1006 [17] Li X-Z, Li Y-L, Yao Y-S, Xie W-D, Zhu J-S. 2020a. Further discussion with Li et al. (2013, 2019)
1007 regarding the "ITS pseudogene hypothesis" for *Ophiocordyceps sinensis*. Mol. Phylogenet. Evol. 146:
1008 106728. (PMID: 31904509 DOI: 10.1016/j.ympev.2019.106728)
1009 [18] Li X-Z, Li Y-L, Yao Y-S, Xie W-D, Zhu J-S, 2022a. Molecular identification of *Ophiocordyceps sinensis*
1010 genotypes and the indiscriminate use of the Latin name for the multiple genotypes and the natural insect-
1011 fungi complex. Am J Biomed Sci. 14(3): 115–135. (doi:10.5099/aj220300115)
1012 [19] Li X-Z, Li Y-L, Yao Y-S, Zhu J-S. 2022b. Differential expressions of the *Hirsutella sinensis* genes and
1013 intraspecific genetic variations among *H. sinensis* strains. Biochem Mol Biol J. 8(2): 58. (DOI:
1014 10.36648/2471-8084.22.8.58)
1015 [20] Li, Y., Hsiang, T., Yang, R.-H., Hu, X.-D., Wang, K., Wang, W.-J., Wang, X.-L., Jiao, L., Yao, Y.-J., 2016a.
1016 Comparison of different sequencing and assembly strategies for a repeat-rich fungal genome,
1017 *Ophiocordyceps sinensis*. J. Microbiol. Methods. 128, 1–6.
1018 (http://dx.doi.org/10.1016/j.mimet.2016.06.025)
1019 [21] Li, Y., Jiang, L., Wang, K., Wu, H.-J., Yang, R.-H., Yan, Y.-J., Bushley, K.E., Hawksworth, D.L., Wu, Z.-
1020 J., Yao, Y.-J., 2020b. RIP mutated ITS genes in populations of *Ophiocordyceps sinensis* and their
1021 implications for molecular systematics. IMA Fungus 11, 18. (PMID: 27343682 DOI:
1022 10.1016/j.mimet.2016.06.025)
1023 [22] Li, Y., Jiao, L., Yao, Y.-J., 2013. Non-concerted ITS evolution in fungi, as revealed from the important
1024 medicinal fungus *Ophiocordyceps sinensis*. Mol. Phylogenetics Evol. 68, 373–379.
1025 (http://dx.doi.org/10.1016/j.ympev.2013.04.010)
1026 [23] Li, Y., Wang, Y.-H., Wang, K., Yang, R.-H., Jiao, L., Yao, Y.-J., 2019. Response to "The multiple
1027 genotypes of *Ophiocordyceps sinensis* and the ITS pseudogene hypothesis". Mol. Phylogenetics Evol.
1028 139, 106522. (PMID: 23618625 DOI: 10.1016/j.ympev.2013.04.010)
1029 [24] Li, Y., Yang, R.-H., Jiang, L., Hu, X.-D., Wu, Z.-J., Yao, Y.-J., 2017. rRNA pseudogenes in filamentous
1030 ascomycetes as revealed by genome data. G3 Genes|Genomes|Genetics. 7, 2695–2703. (PMID: 28637809
1031 doi: 10.1534/g3.117.044016)
1032 [25] Li Y-L, Gao L, Yao Y-S, Li X-Z, Wu Z-M, Lou Z-Q, Xie W-D, Wu J-Y, Zhu J-S. 2022c. Altered GC- and
1033 AT-biased genotypes of *Ophiocordyceps sinensis* in the stromal fertile portions and ascospores of natural

1034 *Cordyceps sinensis*. PLOS ONE (submitted). BIORXIV Preprint 2022-11-12.
1035 (https://www.biorxiv.org/content/10.1101/2022.10.07.511361v1.full.pdf)
1036 [26] Li Y-L, Li X-Z, Yao Y-S, Zhu J-S. 2022d. Differential cooccurrence of multiple genotypes of
1037 *Ophiocordyceps sinensis* in the stromata, stromal fertile portion (ascocarps) and ascospores of natural
1038 *Cordyceps sinensis*. PLOS ONE (in revision). BIORXIV Preprint 2022-6-20.
1039 (https://www.biorxiv.org/content/10.1101/2022.06.20.496766v2.full.pdf)
1040 [27] Li Y-L, Li X-Z, Zhu J-S. 2022e. *Cordyceps sinensis* Fungal Heterokaryons. Chemistry Industry Press Co.,
1041 Beijing, China. (ISBN: 978-7-122-40189-2 http://www.cip.com.cn/Book/Index/48939)
1042 [28] Li, Y.-L., Yao, Y.-S., Xie, W.-D., Zhu, J.-S., 2016b. The molecular heterogeneity of natural *Cordyceps*
1043 *sinensis* with multiple *Ophiocordyceps sinensis* fungi challenges the anamorph-teleomorph connection
1044 hypotheses. Am. J. Biomed. Sci. 8, 123–159. (doi: 10.5099/aj100300217)
1045 [29] Li, Y.-L., Yao, Y.-S., Zhang, Z.-H., Xu, H.-F., Liu, X., Ma, S.-L., Wu, Z.-M., Zhu, J.-S., 2016c. Synergy
1046 of fungal complexes isolated from the intestines of *Hepialus lagii* larvae in increasing infection potency.
1047 J. Fungal Res. 14, 96–112. (DOI: 10.13341/j.jfr.2014.1067)
1048 [30] Liu J, Guo L-N, Li Z-W, Zhou Z, Li Z, Li Q, Bo X-C, Wang S-Q, Wang J-L, Ma S-C, Zheng J, Yang Y.
1049 2020. Genomic analyses reveal evolutionary and geologic context for the plateau fungus *Ophiocordyceps*
1050 *sinensis*. Clin. Med. 15:107–119. (PMID: 33042212 DOI: 10.1186/s13020-020-00365-3)
1051 [31] Liu, Z.-Q., Lin, S., Baker, P., Wu, L.-F., Wang, X.-R., Wu, H., Xu, F., Wang, H.-Y., Brathwaite, M., Zheng,
1052 Y.-G., 2015. Transcriptome sequencing and analysis of the entomopathogenic fungus *Hirsutella sinensis*
1053 isolated from *Ophiocordyceps sinensis*. BMC Genome. 16, 106. (PMID: 25765329 DOI 10.1186/s12864-
1054 015-1269-y)
1055 [32] Liu Z-Y, Yao Y-J, Liang Z-Q, Liu A-Y, Pegler DN, Chase MW. 2001. Molecular evidence for the
1056 anamorph-teleomorph connection in *Cordyceps sinensis*. Mycology Research 105: 827−832.
1057 (https://doi.org/10.1017/S095375620100377X)
1058 [33] Mao, X.-M., Zhao, S.-M., Cao, L., Yan, X., Han, R.-X., 2013. The morphology observation of
1059 *Ophiocordyceps sinensis* from different origins. J. Environ. Entomol. 35, 343–353. (doi:
1060 10.3969/j.issn.1674-0858.2013.03.11)
1061 [34] Ren Y, Wan D-G, Lu X-M, Guo J-L. 2013. The study of scientific name discussion for TCM *Cordyceps*.
1062 Lishizhen Medicine and Materia Medica Research 24: 2211−2212 (Doi: 10.3969/j.issn.1008-
1063 0805.2013.09.075).
1064 [35] Selker EU. 1990. Premeiotic instability of repeated sequences in Neurospora crassa. Annu Rev Genet.
1065 24:579–613. (PMID: 2150906 DOI: 10.1146/annurev.ge.24.120190.003051)
1066 [36] Selker EU. 1997. Epigenetic phenomena in filamentous fungi: useful paradigms or repeat-induced
1067 confusion? Trends Genet. 13:296–301. (PMID: 9260514 DOI: 10.1016/S0168-9525(97)01201-8)

1068 [37] Selker EU. 2002. Repeat-induced gene silencing in fungi. Adv Genet. 46:439–450. (PMID: 11931234
1069 DOI: 10.1016/S0065-2660(02)46016-6)
1070 [38] Selker EU, Stevens JN. 1985. DNA methylation at asymmetric sites is associated with numerous
1071 transition mutations. Proc. Natl. Acad. Sci. USA. 82:8114–8118. (PMID: 2415981 DOI:
1072 10.1073/pnas.82.23.8114)
1073 [39] 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-L. 2020.
1074 A new high-quality draft genome assembly of the Chinese cordyceps *Ophiocordyceps sinensis*. Genome
1075 Biol. Evol. 12(7): 1074−1079. (PMID: 32579174 DOI: 10.1093/gbe/evaa112)
1076 [40] Stensrud, Ø., Hywel-Jones, N,L,, Schumacher, T., 2005. Towards a phylogenetic classification of
1077 Cordyceps: ITS nrDNA sequence data confirm divergent lineages and paraphyly. Mycol Res. 109 (1), 41–
1078 56. (PMID: 15736862 DOI: 10.1017/S095375620400139X)
1079 [41] Stensrud, Ø., Schumacher, T., Shalchian-Tabrizi, K., Svegården, I.B., Kauserud, H., 2007. Accelerated
1080 nrDNA evolution and profound AT bias in the medicinal fungus *Cordyceps sinensis*. Mycol. Res. 111,
1081 409–415. (PMID: 17509850 doi:10.1016/j.mycres.2007.01.015)
1082 [42] Sung, G.-H., Hywel-Jones, N,L,, Sung, J.-M., Luangsa-ard, J.J., Shrestha, B., Spatafora, J.W., 2007.
1083 Phylogenetic classification of *Cordyceps* and the clavicipitaceous fungi. Stud. Mycol. 57, 5–59. (PMID:
1084 18490993 DOI: 10.3114/sim.2007.57.01)
1085 [43] Tan N-Z, Berger JL, Zhang Y, Prolla TA, Weindruch R, Zhao C-S, Bartlett M, Zhu J-S. 2011. The lifespan-
1086 prolonging effect of *Cordyceps sinensis* Cs-4 in normal mice and its molecular mechanisms. FASEB J.
1087 25(1):599.1. (http://www.fasebj.org/cgi/content/meeting_abstract/25/1_MeetingAbstracts/599.1)
1088 [44] Wei, J.-C., Wei, X.-L., Zheng, W., Guo, W., Liu, R., 2016. Species identification and component detection
1089 of *Ophiocordyceps sinensis* cultivated by modern industry. Mycosystema. 35, 404–410. (doi:
1090 10.13346/j.mycosystema.160045)
1091 [45] Wei X-L, Yin X-C, Guo Y-L, Shen N-Y, Wei, J.-C. 2006. Analyses of molecular systematics on *Cordyceps*
1092 *sinensis* and its related taxa. Mycosystema. 25(2): 192–202.
1093 [46] Xia, E.-H., Yang, D.-R., Jiang, J.-J., Zhang, Q.-J., Liu, Y., Liu, Y.-L., Zhang, Y., Zhang, H.-B., Shi, C.,
1094 Tong, Y., Kim, C., Chen, H., Peng, Y.-Q., Yu, Y., Zhang, W., Eichler, E.-E., Gao, L.-Z., 2017. The
1095 caterpillar fungus, *Ophiocordyceps sinensis*, genome provides insights into highland adaptation of fungal
1096 pathogenicity. Sci Rep, 7, 1806. (PMID: 28496210 doi: 10.1038/s41598-017-01869-z)
1097 (www.plantkingdomgdb.com/Ophiocordyceps_sinensis/data/cds/Ophiocordyceps_sinensis_CDS.fas for
1098 the cDNA file)
1099 [47] Xia F, Liu Y, Shen G-R, Guo L-X, Zhou X-W. 2015. Investigation and analysis of microbiological
1100 communities in natural *Ophiocordyceps sinensis*. Can J Microbiol, 61: 104−111. (PMID: 25578897
1101 http://dx.doi.org/10.1139/cjm-2014-0610)

1102 [48] Xiang, L., Li, Y., Zhu, Y.-J., Luo, H.->., Li, C.-F., Xu, X.-L., Sun, C., Song, J.-Y., Shi, L.-C., He, L., Sun,
1103 W., Chen, S.-L., 2014. Transcriptome analysis of the *Ophiocordyceps sinensis* fruiting body reveals
1104 putative genes involved in fruiting body development and cordycepin biosynthesis. Genomics 103, 154–
1105 159. (PMID: 24440419 http://dx.doi.org/10.1016/j.ygeno.2014.01.002)
1106 [49] Xiao, Y-Y., Chen, C., Dong, J.-f., Li, C.-R., Fan, M.-Z. 2011. Morphological observation of ascospores
1107 of *Ophiocordyceps sinensis* and its anamorph in growth process. J Anhui Agricult Univ. 38(4): 587–591.
1108 (http://www.cnki.net/kcms/detail/34.1162.S.20110624.1628.030.html)
1109 [50] Xiao, W., Yang, J.-L., Zhu, P., Cheng, K., He, H., Zhu, H., Wang, Q., 2009. Non-support of species
1110 complex hypothesis of *Cordyceps sinensis* by targeted rDNA-ITS sequence analysis. Mycosystema. 28,
1111 724–730. (DOI: 10.13346/j.mycosystema.2009.05.028)
1112 [51] Yang, J.-L., Xiao, W., He, H., Zhu, H., Wang, S., Cheng, K., Zhu, P., 2008. Molecular phylogenetic
1113 analysis of *Paecilomyces hepiali* and *Cordyceps sinensis*. Acta Pharm. Sin. 43, 421–426. (PMID:
1114 18664207)
1115 [52] Yao, Y.-S., Zhou, Y.-J., Gao, L., Lu, J.-H., Wu, Z.-M., Zhu, J.-S., 2011. Dynamic alterations of the
1116 differential fungal expressions of *Ophiocordyceps sinensis* and its mutant genotypes in stroma and
1117 caterpillar during maturation of natural *Cordyceps sinensis*. J. Fungal Res. 9, 37–49,53. (PMID: 22692321
1118 Doi: 10.3969/j.issn.1672-3538.2011.01.007)
1119 [53] Yao, Y.-S., Zhu, J.-S., 2016. Indiscriminative use of the Latin name for natural *Cordyceps sinensis* insect-
1120 fungi complex and multiple *Ophiocordyceps sinensis* fungi. Chin. J. Chin. Mater. Med. 41 (7), 1316–
1121 1366. (PMID: 28879757 DOI: 10.4268/cjcmm20160733)
1122 [54] Zhang S, Zhang Y-J. 2015. Molecular evolution of three protein-coding genes in the Chinese caterpillar
1123 fungus *Ophiocordyceps sinensis*. Microbiology China 42(8): 1549−1560. (DOI:
1124 10.13344/j.microbiol.china.150279)
1125 [55] Zhang S, Zhang Y-J, Shrestha B, Xu J-P, Wang C-S, Liu X-Z. *Ophiocordyceps sinensis* and *Cordyceps*
1126 *militaris*: research advances, issues and perspectives. Mycosystema 2013; 32: 577−597 (DOI:
1127 10.3969/j.issn.1673-5110.2014.18.061).
1128 [56] Zhang S-W, Cem K, Liu Y, Zhou X-W, Wang C-S. Metatranscriptomics analysis of the fruiting caterpillar
1129 fungus collected from the Qinghai-Tibetan plateau. Scientia Sinica Vitae 2018; 48(5): 562. (doi:
1130 10.1360/N052017-00253)
1131 [57] Zhang Y-J, Li E-W, Wang C-S, Li Y-L, Liu X-Z. *Ophiocordyceps sinensis*, the flagship fungus of China:
1132 terminology, life strategy and ecology. Mycology 2012; 3: 2−10. (DOI: 10.1080/21501203.2011.654354)
1133 [58] Zhang, Y.-J., Sun, B.-D., Zhang, S., Wang, M., Liu, X.-Z., Gong, W.-F., 2010. Mycobiotal investigation
1134 of natural *Ophiocordyceps sinensis* based on culture-dependent investigation. Mycosystema 29, 518−527.
1135 [59] Zhang, Y.-J., Xu, L., Zhang, S., Liu, X.-Z., An, Z., Wang, M., Guo, Y., 2009. Genetic diversity of

1136 *Ophiocordyceps sinensis*, a medicinal fungus endemic to the Tibetan Plateau: implications for its
1137 evolution and conservation. BMC Evolut. Biol. 9, 290. (PMID: 20003548 DOI: 10.1186/1471-2148-9-
1138 290)
1139 [60] Zhao J, Wang N, Chen Y-Q, Li T-H, Qu L-H. 1999. Molecular identification for the asexual stage of
1140 *Cordyceps sinensis*. Acta Scientiarum Naturalium Universitatis Sunyatseni 38: 122−123.
1141 [61] Zhu, J.-S., Gao, L., Li, X.-H., Yao, Y.-S., Zhao, J.-Q., 2010. Maturational alteration of oppositely
1142 orientated rDNA and differential proliferation of GC- and AT-biased genotypes of *Ophiocordyceps*
1143 *sinensis* and *Paecilomyces hepiali* in natural *Cordyceps sinensis*. Am. J. Biomed. Sci. 2, 217–238. (doi:
1144 10.5099/aj100300217)
1145 [62] Zhu J-S, Halpern GM, Jones K. 1998a. The scientific rediscovery of a precious ancient Chinese herbal
1146 regimen: *Cordyceps sinensis*: Part I. J. Altern. Complem. Med. 4(3): 289–303. (PMID: 9764768 DOI:
1147 10.1089/acm.1998.4.3-289)
1148 [63] Zhu J-S, Halpern GM, Jones K. 1998b. The scientific rediscovery of an ancient Chinese herbal medicine:
1149 *Cordyceps sinensis*: Part II. J. Altern. Complem. Med. 4(4): 429–457. (PMID: 9884180 DOI:
1150 10.1089/acm.1998.4.429)
1151 [64] Zhu, J.-S., Li, Y.-L., 2017. A Precious Transitional Chinese Medicine, *Cordyceps sinensis*: Multiple
1152 Heterogeneous *Ophiocordyceps sinensis* in the Insect-Fungi Complex. Lambert Academic Publishing,
1153 Germany. (ISBN-10: 9783330046689, ISBN-13: 978-3330046689)
1154 [65] Zhu, J.-S., Li, Y.-L., Yao, Y.-S., Xie, W.-D., 2018. Molecular evidence not supporting the use of *Hirsutella*
1155 *sinensis* strain EFCC 7287 as the taxonomic standard for multiple *Ophiocordyceps sinensis* fungi. Int. J.
1156 Mol. Biol. 3, 114–115. (DOI: 10.15406/ijmboa.2018.03.00062)
1157 [66] Zhu, J.-S., Li, Y.-L., Yao, Y.-S., Xie, W.-D., 2019. The multiple genotypes of *Ophiocordyceps sinensis*
1158 and the ITS pseudogene hypothesis. Mol. Phylogenetics Evol. 139, 106322. (PMID: 30394296 DOI:
1159 10.1016/j.ympev.2018.10.034)
1160 [67] Zhu, J.-S., Wu, J.-Y., 2015. Genetic heterogeneity of natural *Cordyceps sinensis* with co-existence of
1161 multiple fungi. Chin. J. Cell Biol. 37 (2), 284–298. (DOI: 10.11844/cjcb.2015.02.0284)
1162 [68] Zhu, J.-S., Zhao, J.-G., Gao, L., Li, X.-H., Yao, Y.-S., Zhao, J.-Q., 2012. Dynamically altered expressions
1163 of at least 6 *Ophiocordyceps sinensis* mutants in the stroma of *Cordyceps sinensis*. J. Fungal Res. 10, 100–
1164 112. (DOI: 10.13341/j.jfr.2012.02.007)
1165

Multiple GC-biased repetitive ITS copies in the *Hirsutella sinensis* genome are not generated by RIP mutagenesis involving transition point mutations

Xiu-Zhang Li[a], Yu-Ling Li[a], Ya-Nan Wang[b], Jia-Shi Zhu[a,c,*]

Table 1. *GC contents of the repetitive ITS1-5.8S-ITS2 sequences from the genome assemblies of 2 H. sinensis strains and percent similarities of the repetitive ITS sequences to the sequences of GC-biased Genotype #1 (AB067721) and AT-biased Genotypes #4–6 and #15–17 of O. sinensis.*

Genome segment	Sequence range & direction	GC content	Percent similarity of the repetitive sequence	
			vs. Genotype #1 AB067721	*vs*. AT-biased Genotypes #4–6 and #15–17
JAAVMX010000002	18688917→18689407	64.8%	100%	86.7% – 89.9%
	18702095→18702586	64.5%	97.4%	82.8% – 88.0%
JAAVMX010000008	13823→14313	64.8%	100%	85.5% – 89.9%
	1199→1687	64.8%	99.2%	85.2% – 89.3%
JAAVMX010000017	9147←9637	64.8%	100%	85.5% – 89.9%
	21791←22281	64.8%	100%	85.5% – 89.9%
	34435←34925	64.8%	100%	85.5% – 89.9%
	47079←47569	64.8%	100%	85.5% – 89.9%
JAAVMX010000018	13381→13871	64.8%	100%	85.5% – 89.9%
	26076→26566	64.8%	100%	85.5% – 89.9%
	38771→39261	64.8%	100%	85.5% – 89.9%
	51467→51958	64.6%	98.8%	85.5% – 88.8%
	700→1186	64.5%	97.0%	82.6% – 87.6%
JAAVMX010000019	19404→19894	64.8%	100%	85.5% – 89.9%
	32048→32538	64.8%	100%	85.5% – 89.9%
	6233→6733	65.3%	94.5%	81.3% – 85.5%
	44729→45251	63.3%	90.8%	80.1% – 84.6%
NGJJ01000573	13133←13623	64.8%	99.4%	85.2% – 89.6%
	14624←15113	64.7%	99.4%	84.8% – 89.6%
	1←260 (ITS1 and partial 5.8S)	56.5%	98.5%	83.5% – 88.3%
NGJJ01000582	10703←11192	64.7%	99.8%	85.5% – 89.8%
	24068←24190 (only partial ITS2)	73.2%	98.1%	87.7% – 89.1%
NGJJ01000796	1514→2004	64.8%	100%	85.5% – 89.9%
NGJJ01000798	1177→1665	64.6%	99.6%	84.8% – 89.6%
NGJJ01000799	3295→3785	64.8%	100%	85.5% – 89.9%
	16666→17156	64.8%	100%	85.5% – 89.9%

10 ***Table 2.*** *Percent similarities between nrDNA ITS1-5.8S-ITS2 sequences of 5 genome assemblies of*
11 *Genotype #1 H. sinensis and 17 O. sinensis genotype sequences. (Modified from Table 1 of Li* et al.
12 *2022d).*

GT	GenBank Accession #	% similarity between the *H. sinensis* genomic ITS segments and 17 *O. sinensis* genotypes				
		ANOV01021709 838→1504	LKHE01000582 2074→2740	LWBQ01000008 991739→992405	JAAVMX010000017 9029←9695	NGJJ01000799 3237→3903
#1	AB067721	99.7%	100%	99.6%	100%	100%
#2	MG770309	94.4%	94.7%	94.7%	94.7%	94.7%
#3	HM595984	95.9%	96.1%	95.6%	95.3%	95.3%
#7	AJ488254	94.7%	95.0%	95.3%	95.0%	95.0%
#8	GU246286	89.5%	89.5%	89.2%	89.5%	89.5%
#9	GU246288	95.0%	95.2%	94.6%	95.2%	95.2%
#10	GU246287	83.0%	83.2%	83.0%	83.2%	83.2%
#11	JQ695935	78.9%	79.1%	78.5%	79.1%	79.1%
#12	GU246296	94.3%	94.5%	93.9%	94.5%	94.5%
#13	KT339190	87.5%	87.5%	87.2%	87.5%	87.5%
#14	KT339178	89.3%	89.6%	89.1%	89.6%	89.6%
#4	AB067744	89.1%	89.4%	89.3%	89.4%	89.4%
#5	AB067740	86.3%	86.7%	86.2%	86.7%	86.7%
#6	KJ720572	85.5%	85.5%	85.5%	85.5%	85.5%
#15	KT232017	89.6%	89.9%	89.4%	89.9%	89.9%
#16	KT232019	87.4%	87.7%	87.6%	87.7%	87.7%
#17	KT232010	87.7%	88.1%	87.5%	88.1%	88.1%

13 *Note: "GT" represents genotype. GC-biased genotypes are listed in blue in the upper panel. AT-*
14 *biased genotypes are listed in red in the lower panel.*
15
16 ***Table 3.*** *Types of mutations in the variable and slightly less variable genomic repetitive ITS*
17 *sequences compared with GC-biased Genotype #1 (AB067721).*

	Sequence range & direction	% similarity	Number of mutant alleles			Ins./Del./Transv. *vs*. Transit.
			Ins./Del.	Transv.	Transit.	
JAAVMX010000002	18702095→18702586	97.4%	13	0	0	13 : 0
JAAVMX010000018	700→1186	97.0%	10	5	0	15 : 0
JAAVMX010000019	6233→6733	94.5%	22	4	1	26 : 1
	44729→45251	90.8%	32	11	5	43 : 5

18 *Note: "Ins./Del." indicates insertion/deletion point mutations. "Transv." denotes Transversion point*
19 *mutations. "Transit." refers to transition point mutations.*
20

```
GT#1  AB067721  59        TCGAGTCACCACTCCC AAACCCCCTGCGAACACCACAGCAGTTGCCTCGGCGGGACCGCCCCGGCG CCCCAGGGCCCG 136
GT#3  FJ654148  1         CT----TG-------- ---A----C-T---------------------------------------- ------------ 83
JAAVMX010000019 19404     ---------------- ---------------------------------------------------- ------------ 19481
JAAVMX010000019 32048     ---------------- ---------------------------------------------------- ------------ 32125
JAAVMX010000002 18702095  ---------------- ----------------------- ---------------------------T------------ 18702172
JAAVMX010000018 700       --------- ------ ---------------------------------------T------------ ------------T 776
JAAVMX010000019 6233      ----------------C--------------------------------------------------- ------------ 6311
JAAVMX010000019 44729     ---------------- ---------------------------------------------------- ------------ 44806
GT#4  AB067744  55        ------T--------T --------------T--T-------------------------------- ---T-------- 132
GT#5  AB067740  59        ------T--T----TT ---------ATA--T--T-T-----------TA------------TA--- ---T-A------ 136
GT#6  KJ720572  11                         -------A----T--T-T-----------TA------------TA--- --TT-A------ 70
GT#15 KT232017  37        ------T--------T --------------T--T------------------------------- ---T-------- 114
GT#16 KT232019  37        ------T--T-----T --------------T--T-T-----------TA------------TA--- ---T-A------ 114
GT#17 KT232010  37        ------T--T-----T --------------T--T-T-----------TA------------TA--- ---T-A------ 114

GT#1  AB067721  137       GACCAGG GCGCCC GCCGGAGGACCCCCAGAC CCTCCTGTCGCA GTGGCATCT CTCAGTCAAGAA GCAAGCAAAT 210
GT#3  FJ654148  84        ------- ------ ------------------ T----------- -----G--- ------------ ---------- 157
JAAVMX010000019 19482     ------- ------ ------------------ ------------ --------- ------------ ---------- 19555
JAAVMX010000019 32126     ------- ------ ------------------ ------------ --------- ------------ ---------- 32199
JAAVMX010000002 18702173  ------- ------A------------------ -------------A--------- ---- ------- ---------- 18702247
JAAVMX010000018 777       ------- ------ ------------------ ------------ --------- ---C-------- ---------- 850
JAAVMX010000019 6312      ------- ------ ------T-----------C------------ ---------C------------G---------- 6388
JAAVMX010000019 44807     -------C------ ------------------ ------------ --------- ------------ ---------- 44881
GT#4  AB067744  133       ---T--- ------ ------------TT---- ------A---T- --A------ --T---T----- ---------- 206
GT#5  AB067740  137       ---T-AA ------ ------------TT---- ------A-TAT- --A-T---- --T---T----- ---------- 213
GT#6  KJ720572  71        ---T--- ------ ------------TT---- ------A---T- --A-T---- --T---T----- ---------- 182
GT#15 KT232017  115       ---T--- ------ ------------TT---- ------A----- --A------ --T---T----- ---------- 188
GT#16 KT232019  115       ---T-AA ------ ------------TT---- ------A----- --A-T---- --T---T----- ---------- 188
GT#17 KT232010  115       ---T-AA ------ ------------TT---- ------A-TAT- --A-T---- --T---T----- ---------- 188

GT#1  AB067721  211       GAATCAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATC GATG AAGAACGCAGCGAAATGCGATAAGTAATGTGAA 288
GT#3  FJ654148  158       ---------------------------------------- ---- ---------------------------------- 235
JAAVMX010000019 19556     ---------------------------------------- ---- ---------------------------------- 19633
JAAVMX010000019 32200     ---------------------------------------- ---- ---------------------------------- 32277
JAAVMX010000002 18702248  -- ------------------------------------C---- ---------------------------------- 18702325
JAAVMX010000018 851       ---------------------------------------- ----A----------------------- --------- 928
JAAVMX010000019 6389      -------------- ---CAC-------------------- ---- --------------- --------- ------- 6463
JAAVMX010000019 44882     ---------------------------------------- ---- ---------------------------------- 44959
GT#4  AB067744  207       A---T--------T--T------------A---T-A----T A--A ----------------A-------------A-- 284
GT#5  AB067740  214       A---T----------T---------------A-----A----T A--A --A----------------A---------A-A-- 288
GT#6  KJ720572  183       A---T--------T---------------A-----A----- ---A -------T--T-----A-----------A-A-- 222
GT#15 KT232017  189       ----T--------T--T------------A---T-A----T A--A ----------------A-------------A-- 266
GT#16 KT232019  189       ----T--------T---------------A-----A----- ---A --A----------------A---------A-A-- 266
GT#17 KT232010  189       A---T--------T---------------A-----A----- ---A --A----------------A---------A-A-- 266

GT#1  AB067721  289       TTGCAGAATTCAGTGAACCA TCGAATCTTT GAACG CACATTG CGCCCGCCAGCACTCTGGCGGGCATGC  CTGTC 362
GT#3  FJ654148  236       -------------------- ---------- ----- ------- ---------------------------  ----- 309
JAAVMX010000019 19634     -------------------- ---------- ----- ------- ---------------------------  ----- 19707
JAAVMX010000019 32278     -------------------- ---------- ----- ------- ---------------------------  ----- 32351
JAAVMX010000002 18702326  -------------------- ----------T----- ----- - ------- ---------------------  ----- 18702398
JAAVMX010000018 929       ------------------- C---------- -----C-------C--------------------- -------  ----- 1000
JAAVMX010000019 6464      --------------------C---------- ----- -------C---------------------------  ----- 6539
JAAVMX010000019 44960     -------------------- ---------- ----- ------- ---------------------------CT----- 45035
GT#4  AB067744  285       ---T------T---A---T- ---------- A---- T-T-C-A -------T--------A----------  ----- 358
GT#5  AB067740  289       --A-------T---A---T- ---------- A---- --T---A -------T--------A----------  --A-- 362
GT#6  KJ720572  223       --AT------T---A----- ---------- A---- T-T---A -------T--------A----------  ----- 296
GT#15 KT232017  267       ---T------T---A---T- ---------- A---- T-T---A -------T--------A----------  ----- 340
GT#16 KT232019  267       --A-------T---A---T- ---------- A---- T-T---A -------T--------A----------  ----- 340
GT#17 KT232010  267       --A-------T---A---T- ---------- A---- --T---- -------T--------A----------  --A-- 340

GT#1  AB067721  363       CGAGCGTCATCTCAACCCTCGAG CCCCCCGCCTCGCGG  CGGCGGGGCCCGGCCTT GGGGGTCACGGCC CCGCGCC 437
GT#3  FJ654148  310       ------------------------ --------------  ----------------- ------------- ------- 384
JAAVMX010000019 19708     ------------------------  --------------  ----------------- ------------- ------- 19781
JAAVMX010000019 32352     ------------------------  --------------  ----------------- ------------- ------- 32425
JAAVMX010000002 18702399  ------------------------ ----- ---------  ----------------- ------------- ------- 18702472
JAAVMX010000018 1001      ------------------------ ------ --------  -----------------G------------- ---- -- 1074
JAAVMX010000019 6540      ------------------------  --------------  ------------- ---- -------------C------- 6613
JAAVMX010000019 45036     ------------------------C----------A----TA--CG---C--------- ------TC-----AA------ 45119
GT#4  AB067744  359       -------T-C--T------TA-- ---------------  ----------------- A------------ --A---- 432
GT#5  AB067740  363       -------T----T---------- ---------------  ----------T------- A-----T------ ------- 436
GT#6  KJ720572  297       -------T-C--T----                                                                313
GT#15 KT232017  341       -------T----T---------- ---------------  ----------T------- A-----T------ ------- 414
GT#16 KT232019  341       -------T-C--T------TA-- ---------------  ----------------- A------------ --A---- 414
GT#17 KT232010  341       -------T----T---------- ---------------  ----------T------- A-----T------ ------- 414

GT#1  AB067721  438       G   CCCCCTAAACGC AG       TGGCGACC  CCGCCGCGGCTCC CCTGCGCAGTAGCT  CG  CTGAGAACCT 499
GT#3  FJ654148  385       -   ----------T-- --       --------  ------------- ---------------  --  -C---G---C 446
JAAVMX010000019 19782     -   ------------ --       --------  ------------- ---------------  --  ---------- 19844
JAAVMX010000019 32426     -   ------------ --       --------  ------------- ---------------  --  ---------- 32488
JAAVMX010000002 18702473  -   ------------ --       --------  ------------- ---------------  --  ---------- 18702534
JAAVMX010000018 1075      -   ------------ --       --------  ------------- ---------------  --  ---------- 1136
JAAVMX010000019 6614      -    -----------G--G      --------GC------------- ---------------  --GCT--------- 6680
JAAVMX010000019 45120     -TCC-T---GT----- --CCTAAGTG-------  ----GA-------A----T-A-------CA--  ---------- 45189
GT#4  AB067744  433       -   ------------ --       -A-T----  ------------- ---A----A-----  TA  --A------- 495
GT#5  AB067740  437       -   ----T------T --       -A------  ------------- ---A--TG------  -A  -C-------- 529
GT#6  KJ720572  313
GT#15 KT232017  415       -   ------------T --       -A------  ------------- ---A--T-A-----  -A  -C-------- 477
GT#16 KT232019  415       -   ------------ --       -A------  ------------- ---A----A-----  TA  --A------- 477
GT#17 KT232010  415       -   -----------T --       -A------  ------------- ---A--T-A-----  -A  -C-------- 477

GT#1  AB067721  500        CGCACCGGGAGCGCGGA     GGCGG TCACGCC  TGG    AA ACCACC   ACACCCTCCA              549
GT#3  FJ654148  447        --------A-------     ----- -------  ---    -- ------   ---     ---             492
JAAVMX010000019 19845      ------------------     ----- ----- -  ---    -- ------   ----------             19894
JAAVMX010000019 32489      ------------------     ----- ----- -  ---    -- ------   ----------             32537
JAAVMX010000002 18702535   ------------------     ----- -------GTGT-    -- ------   ----------             18702586
JAAVMX010000018 1137       ------------------     ----- ----- -  ---    -- ------   ----------             1186
JAAVMX010000019 6681      C------------------     ----- ----- -  ---    -- ------ACA----------             6733
JAAVMX010000019 45190      ------------------CCCAT-----C-------  ---GATG--C------ACC----------             45251
GT#4  AB067744  496        -----TA----------     ----- -------  --A    -- -----T   -T------              543
GT#5  AB067740  530        --------A--------     ----- -------  --A    -- -----T   --G--                 544
GT#6  KJ720572  313
GT#15 KT232017  478        --------A--------     ----- -------  --A    -- -----T   --G--                 522
GT#16 KT232019  478        -----TA----------     ----- -------  --A    -- -----T   -T------              525
GT#17 KT232010  478        --------A--------     ----- -------  --A    -- -----T   --G--                 522
```

Figure 1. ITS sequence alignment of AB067721, repetitive ITS copies within genome JAAVMX000000000 of H. sinensis Strain IOZ07, and AT-biased genotypes of O. sinensis.

The ITS sequences contain complete or partial ITS1-5.8S-ITS2 nrDNA segments. "GT" denotes the O. sinensis genotype. The underlined sequence in black represents the 5.8S gene of GC-biased Genotype #1 H. sinensis. AB067721 is the ITS sequence of GC-biased Genotype #1 H. sinensis. Genome assemblies JAAVMX010000002, JAAVMX010000018 and JAAVMX010000019 were obtained from H. sinensis Strain IOZ07 (Shu et al. *2020). One copy each within JAAVMX010000002 and JAAVMX010000018 indicated in* ***green*** *shows 97.4% or 97.0% similarity to AB067721. JAAVMX010000019 contains 4 repetitive ITS copies, including 2 copies in* ***black*** *(19404→19894 and 32048→32537), which are 100% identical to AB067721, and 2 other copies in* ***blue*** *(6233→6733 and 44729→45251), with 94.5% and 90.8% similarities to AB067721. Sequences in* ***red*** *(AB067744, AB067740, KJ720572, KT232017, KT232019, and KT232010) represent AT-biased Genotypes #4–6 and #15–17 of O. sinensis, respectively. The underlined "GAATTC" sequences in* ***purple*** *represent the Eco*RI *endonuclease cleavage sites in the sequences of GC-biased Genotype #1 and GC-biased genome assembly JAAVMX000000000. Eco*RI *endonuclease cleavage sites are absent in AT-biased ITS sequences due to a single-base C-to-T (cytosine→threonine) transition. Hyphens indicate identical bases, and spaces denote unmatched sequence gaps.*

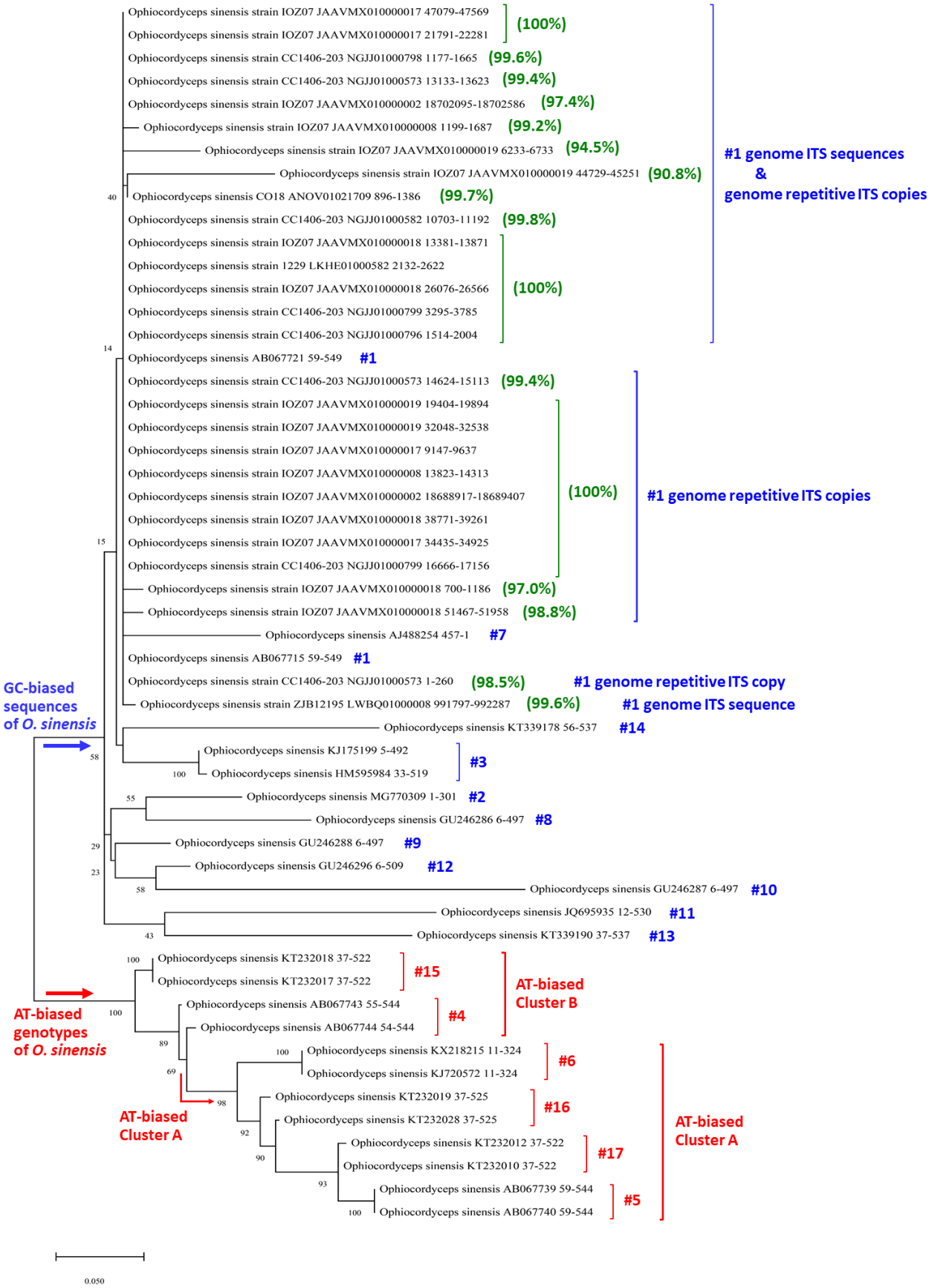

Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 47079-47569
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 21791-22281
(100%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000798 1177-1665 (99.6%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000573 13133-13623 (99.4%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000002 18702095-18702586 (97.4%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000008 1199-1687 (99.2%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 6233-6733 (94.5%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 44729-45251 (90.8%)
40
Ophiocordyceps sinensis CO18 ANOV01021709 896-1386 (99.7%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000582 10703-11192 (99.8%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 13381-13871
Ophiocordyceps sinensis strain 1229 LKHE01000582 2132-2622
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 26076-26566
(100%)
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000799 3295-3785
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000796 1514-2004
#1 genome ITS sequences
&
genome repetitive ITS copies
14
Ophiocordyceps sinensis AB067721 59-549 #1
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000573 14624-15113 (99.4%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 19404-19894
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000019 32048-32538
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 9147-9637
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000008 13823-14313
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000002 18688917-18689407
(100%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 38771-39261
15
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000017 34435-34925
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000799 16666-17156
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 700-1186 (97.0%)
Ophiocordyceps sinensis strain IOZ07 JAAVMX010000018 51467-51958 (98.8%)
#1 genome repetitive ITS copies
Ophiocordyceps sinensis AJ488254 457-1 #7
Ophiocordyceps sinensis AB067715 59-549 #1
Ophiocordyceps sinensis strain CC1406-203 NGJJ01000573 1-260 (98.5%) #1 genome repetitive ITS copy
Ophiocordyceps sinensis strain ZJB12195 LWBQ01000008 991797-992287 (99.6%) #1 genome ITS sequence
GC-biased sequences of O. sinensis
Ophiocordyceps sinensis KT339178 56-537 #14
58
Ophiocordyceps sinensis KJ175199 5-492
100
Ophiocordyceps sinensis HM595984 33-519
#3
55
Ophiocordyceps sinensis MG770309 1-301 #2
Ophiocordyceps sinensis GU246286 6-497 #8
29
Ophiocordyceps sinensis GU246288 6-497 #9
23
Ophiocordyceps sinensis GU246296 6-509 #12
58
Ophiocordyceps sinensis GU246287 6-497 #10
Ophiocordyceps sinensis JQ695935 12-530 #11
43
Ophiocordyceps sinensis KT339190 37-537 #13
100
Ophiocordyceps sinensis KT232018 37-522
Ophiocordyceps sinensis KT232017 37-522
#15
AT-biased Cluster B
AT-biased genotypes of O. sinensis
100
Ophiocordyceps sinensis AB067743 55-544
Ophiocordyceps sinensis AB067744 54-544
#4
89
100
Ophiocordyceps sinensis KX218215 11-324
Ophiocordyceps sinensis KJ720572 11-324
#6
69
98
Ophiocordyceps sinensis KT232019 37-525
Ophiocordyceps sinensis KT232028 37-525
#16
AT-biased Cluster A
92
90
Ophiocordyceps sinensis KT232012 37-522
Ophiocordyceps sinensis KT232010 37-522
#17
93
Ophiocordyceps sinensis AB067739 59-544
100
Ophiocordyceps sinensis AB067740 59-544
#5
AT-biased Cluster A
0.050

Figure 2. A Bayesian majority rule consensus phylogenetic tree.

Twenty-eight ITS segments of the genome assemblies (ANOV01021709, LKHE01000582, LWBQ01000008, JAAVMX010000002, JAAVMX010000008, JAAVMX0100000017, JAAVMX0100000018, JAAVMX010000019, NGJJ01000573, NGJJ01000582, NGJJ01000796, NGJJ01000798, and NGJJ01000799) of H. sinensis strains (Co18, 1229, ZJB12195, IOZ07, and CC1406-203, respectively) and 25 ITS sequences of GC-biased Genotypes #1–3 and #7–14 (in blue alongside the tree) and AT-biased Genotypes #4–6 and #15–17 of O. sinensis (in red alongside the tree) were analyzed phylogenetically using MrBayes v3.2.7 software (cf. Section II-4). Genome assemblies JAAVMX000000000 and NGJJ00000000 contain multiple repetitive ITS copies (cf. Table 1). The percent similarities of the genomic ITS sequences with the representative Genotype #1 sequence (AB067721) are shown in green alongside the tree.

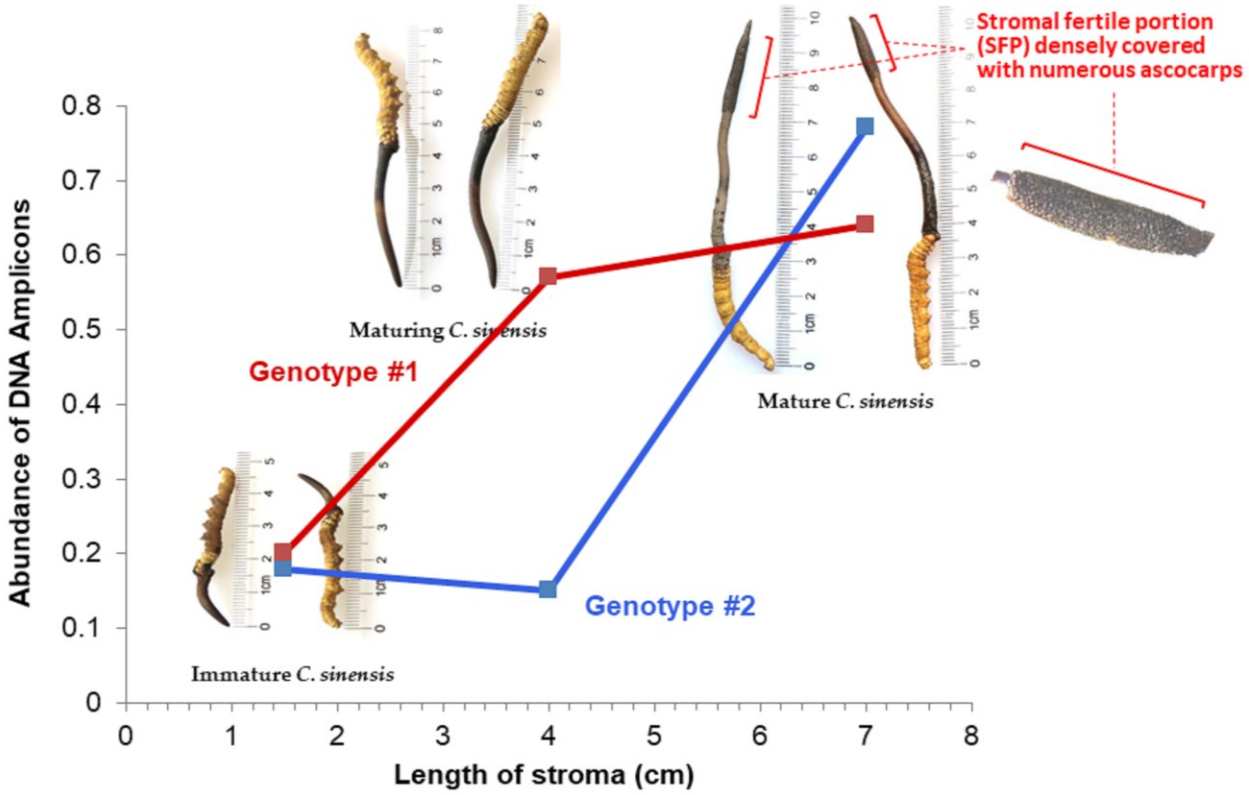


Figure 3. Dynamic alterations in the abundance of amplicons of Genotypes #1 and #2 of O. sinensis in the stromata of C. sinensis during maturation *(modified from Figure 6 of Zhu* et al. *2012).*

Immature C. sinensis specimens had very short stromata of 1.5 cm on average. Maturing C. sinensis specimens had stromata of 4.0 cm on average, without the formation of an expanded fertile portion close to the stromal tip. Mature C. sinensis specimens had long stromata of 7.0 cm on average and showed the formation of an expanded stromal fertile portion (SFP in red), which was densely covered with numerous ascocarps.

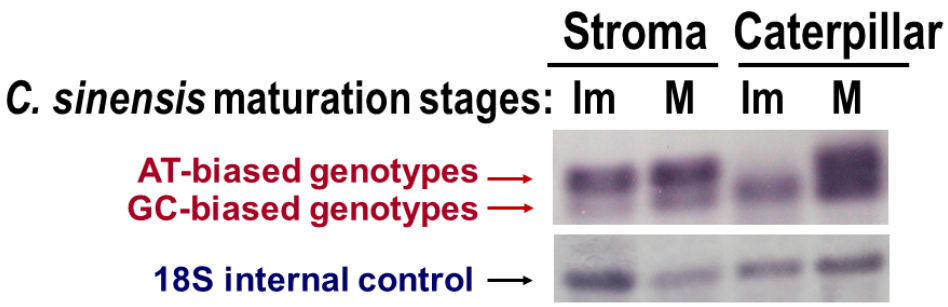


Figure 4. Southern blot of C. sinensis nrDNA in the stroma and caterpillar body of natural C. sinensis during maturation. *[Reproduced with permission from AJBMS (www.nwpii.com/ajbms) Am J Biomed Sci 2010; 2(3): 217-238] [Zhu* et al. *2010].*

Caterpillar refers to caterpillar body. Genomic DNA templates were isolated from the stromata or caterpillar bodies of immature (Im) or mature (M) C. sinensis specimens collected from Sichuan Province in China and prepared using the restriction enzymes EcoRI, DraI and AvaI. ***The upper panel*** *was probed with an H. sinensis-specific probe.* ***The lower panel*** *was probed with a nonspecific 18S internal control probe.*

Figure 5. Alignment of the query sequence in green (the 5.8S gene of Genotype #1 H. sinensis AB067721) and the 75 of 1894 overlapping (pink bars) shotgun reads (GenBank SRR5428527) of mature C. sinensis.

68 ***Table 4.*** *Percent similarity between the unassembled metatranscriptome shotgun reads (SRA:*
69 *SRR5428527) of natural C. sinensis and the 5.8S genes of multiple genotypes of O. sinensis and*
70 *Paecilomyces hepiali, an AB067719-type fungus (Group-E) and Tolypocladium sinensis (Group D).*

Geno-type	GenBank Accession #	% similarity between 5.8S genes and *C. sinensis* shotgun sequences					
		SRR5428527.17546327.1	SRR5428527.17554054.1	SRR5428527.17565575.2	SRR5428527.17679995.2	SRR5428527.17911498.2	SRR5428527.17929569.2
#1	AB067721	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
#2	MG770309	96%	96%	**100%**	96%	**98%**	96%
#3	HM595984	**100%**	**99%**	**99%**	**100%**	**99%**	**100%**
#7	AJ488254	**98%**	**98%**	**99%**	**98%**	**98%**	**98%**
#8	GU246286	94%	96%	95%	94%	**97%**	94%
#9	GU246288	**99%**	**99%**	**99%**	**100%**	**99%**	**100%**
#10	GU246287	96%	94%	93%	95%	93%	**97%**
#11	JQ695935	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
#12	GU246296	**99%**	**99%**	**99%**	**99%**	**99%**	**99%**
#13	KT339190	**98%**	93%	93%	**98%**	93%	**98%**
#14	KT339178	**98%**	93%	93%	**98%**	93%	**98%**
#4	AB067744	86%	87%	87%	83%	87%	83%
#5	AB067740	85%	87%	87%	84%	87%	84%
#6	KJ720572	85%	85%	87%	84%	86%	84%
#15	KT232017	85%	88%	88%	84%	88%	84%
#16	KT232019	86%	86%	87%	85%	87%	85%
#17	KT232010	87%	88%	88%	87%	88%	87%
P.h.	EF555097	**99%**	**98%**	**97%**	**99%**	**98%**	**99%**
	AB067719	**98%**	93%	93%	**98%**	93%	**98%**
T.s.	AB067720	**98%**	**97%**	**96%**	**98%**	**97%**	**98%**

71 *Note: GC-biased genotypes are listed in blue, and AT-biased genotypes are listed in red. GenBank*
72 *sequence ID SRA: SRR5428527 fragments (17546327.1, 17554054.1, 17565575.2, 17679995.2,*
73 *17911498.2, and 17929569.2) are the unassembled transcriptome shotgun sequences obtained from*
74 *the mature C. sinensis specimen collected from Deqin County of Yunnan Province by Xia* et al. *(2017).*
75 *P.h. refers to Paecilomyces hepiali; AB067719 refers to an AB067719-type Group-E fungus; T.s.*
76 *refers to Group-D Tolypocladium sinensis (Stensrud* et al. *2007). Highly homologous percent*
77 *similarities (97%–100%) are bolded in* ***green****.*
78

```
GT1 AB067721 1     TAGAGGAAGTAAAAGTCGTAACAAGGTCTCCGTTGGTGAACCAGCGGAGGGATCATTATC  60
GT4 AB067744 1         -------------------------------A--A---T-----------T------  56
GT5 AB067740 1     --A------------------T------------A--A---T-----------T------  60

GT1 AB067721 61    GAGTCACCACTCCCAAACCCCCTGCGAACACCACAGCAGTTGCCTCGGCGGGACCGCCCC  120
GT4 AB067744 57    ------------C--------------A--C-----------------------------  116
GT5 AB067740 61    ----T--T----TT---------ATA--T--T-T-----------TA------------T  120

GT1 AB067721 121   GGCGCCCCAGGGCCCGGACCAGGGCGCCCGCCGGAGGACCCCCAGACCCTCCTGTCGCAG  180
GT4 AB067744 117   -------T-----------T---------------------TT----------A---T--  176
GT5 AB067740 121   A------T-A---------T-AA------------------TT----------A-TAT--  180

GT1 AB067721 181   TGGCATCTCTCAGTCAAGAAGCAAGCAAATGAATCAAAACTTTCAACAACGGATCTCTTG  240
GT4 AB067744 177   -A--------T---T---------------A---T-A-ACTTTTAATAACGGATCTCTTA  236
GT5 AB067740 181   -A-T------T---T---------------A---T---ACTTTTAACAACGGATCTCTTA  240
Primer 5.8S-F                                            ACTTTTAACAACGGATCTCTT

GT1 AB067721 241   GTTCTGGCATCGATGAAGAACGCAGCGAAATGCGATAAGTAATGTGAATTGCAGAATTCA  300
GT4 AB067744 237   ---T-A----TA--A----------------A-------------A-----T------T-  296
GT5 AB067740 241   -----A----TA--A--A---------------A---------A-A----A-------T-  300

GT1 AB067721 301   GTGAACCATCGAATCTTTGAACGCACATTGCGCCCGCCAGCACTCTGGCGGGCATGCCTG  360
GT4 AB067744 297   --A---T-----------A----T-T-C-A-------T--------A------ATGCCTG  356
GT5 AB067740 301   --A---T-----------A------T---A-------T--------A------ATGCCTA  360
Primer 5.8S-R                                                         ATGCTTA

GT1 AB067721 361   TCCGAGCGTCATCTCAACCCTCGAGCCCCCCGCCTCGCGGCGGCGGGGCCCGGCCTTGGG  420
GT4 AB067744 357   TCCGAGCGTTACCTT------TA----------------------------------A--  416
GT5 AB067740 361   TCCGAGCGTTATCTT--------------------------------T-------A--  420
Primer 5.8S-R      TCCGAGCGTTATCTT

GT1 AB067721 421   GGTCACGGCCCCGCGCCGCCCCCTAAACGCAGTGGCGACCCCGCCGCGGCTCCCCTGCGC  480
GT4 AB067744 417   ------------A--------------------A-T--------------------A---  476
GT5 AB067740 421   ---T------------------T------T---A----------------------A--T  480

GT1 AB067721 481   AGTAGCTCGCTGAGAACCTCGCACCGGGAGCGCGGAGGCGGTCACGCCGTGAAACCACCA  540
GT4 AB067744 477   -A-----TA--A------------TA------------------------A-------T-  536
GT5 AB067740 481   G-------A-C----------------A----------------------A-------T-  540

GT1 AB067721 541   CACCCTCCAGTTGACCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCAATAA  600
GT4 AB067744 537   T------TG---A---------T-----------------A-------------------  596
GT5 AB067740 541   -G--T-TT----A---------T-----------------A-------------T-----  600

GT1 AB067721 601   GCGGAGGAAAAGAAACCAACAGGGATTGCCCCAGTAACGGCGAGTGAAGCGGCAACAGCT  660
GT4 AB067744 597   --------                                                       604
GT5 AB067740 601   ----------------T--T----------TT-------------A------T--T----  660
```

Figure 6. ITS sequence alignment of AB067721 of GC-biased Genotype #1, AB067744 and AB067740 of AT-biased Genotypes #4–5 of O. sinensis and the 5.8S-F/R primers.

"GT1", "GT4" and "GT5" represent Genotypes #1 and #4–5 of O. sinensis, respectively. The sequences in blue are 18S and 28S genes (5' and 3' end regions, respectively). Those in pink represent sequences of 5.8S genes. Those in black are ITS1 between the 18S and 5.8S genes and ITS2 between the 5.8S and 28S genes. The sequences underlined in green are for primers 5.8S-F/R, while primer 5.8S-R was converted to the sense chain sequence. Hyphens indicate identical bases.