



Density-weighted Algorithms for Similarity Computation and Cluster Tree Construction in the RAPD Analysis of Natural *Cordyceps sinensis*

Luqun Ni ^{1#}, Yi-Sang Yao ^{2#}, Ling Gao ², Zi-Mei Wu ², Ning-Zhi Tan ², Jia-Shi Zhu ^{3,4,*}

¹ Department of Mechanical and Aerospace Engineering, University of California San Diego, La Jolla, CA, USA

² Pharmanex Beijing Pharmacology Center, Beijing 100088, China

³ Department of Applied Biology and Chemistry Technology, Hong Kong Polytechnic University, Hong Kong

⁴ NS Center for Anti-aging Research, Provo, UT 84601, USA

These authors (LN and YSY) contributed equally to this work

*Corresponding author:

Dr. J-S Zhu

NS Center for Anti-aging Research

Provo, UT 84601

USA

Tel: +1(858) 705-3790

Fax: +1(858) 777-5435

E-mail: zhujosh@gmail.com

Received: 6 March 2014; | Revised: 25 April 2014; | Accepted: 16 May 2014

Abstract

Objective: The goal of this study was to develop and validate density-weighted arithmetic methods for the analysis of the dynamic maturational changes in random amplified polymorphic DNA (RAPD) polymorphisms in *Cordyceps sinensis* containing multiple fungi. **Methods:** Ten random primers were used for PCR amplification to monitor changes in the RAPD molecular marker polymorphisms in caterpillar body, stroma and ascocarp portion samples of *C. sinensis* collected at different stages of maturation. We compared (1) the density-unweighted Nei-Li similarity equation and new similarity equations considering the densities of DNA bands and (2) the density-unweighted and weighted arithmetic methods for cluster analysis. **Results:** The algorithm using the Nei-Li similarity equation did not account for the differences in the density of the matched randomly amplified amplicons, whereas the new similarity equations were capable of integrating different amplicon densities into the similarity computation. With improvements in the similarity computation and cluster construction, the new methods revealed dynamic maturational changes in the RAPD polymorphisms of molecular markers in *C. sinensis* caterpillar bodies and stromata. The polymorphism analysis revealed similarities of 74%-88% between the RAPD polymorphisms of the ascocarp portion versus premature and mature *C. sinensis* stromata, but low similarities of <70% versus the caterpillar bodies and maturing stromata. Apparent dissimilarities (similarity <66%) were also found between *Hirsutella sinensis* and *C. sinensis* samples. **Conclusions:** These new algorithmic methods represent an advance in similarity computations when comparing polymorphisms of RAPD molecular markers in *C. sinensis* and provide accurate analytic tools to capture the molecular information related to the dynamic

changes in differential expression of *C. sinensis* intrinsic fungi during maturation. The apparent dissimilarity in RAPD polymorphisms between *H. sinensis* and *C. sinensis* suggests differences of mycological background and challenges the hypothesis of the anamorph-teleomorph connection between *H. sinensis* and *Ophiocordyceps sinensis*.

Keywords: RAPD (Random Amplified Polymorphic DNA) polymorphism of molecular markers; density-weighted similarity computation; density-weighted cluster construction; *Cordyceps sinensis*; *Ophiocordyceps sinensis*; *Hirsutella sinensis*.

1. Introduction

Natural *Cordyceps sinensis* (Berk.) Sacc.¹, which has been used for centuries as a precious medicinal herb in China and other Asian countries, has a broad spectrum of clinical benefits, including anti-aging effects [1–5]. Using microcosmic molecular techniques, such as molecular cloning, rDNA ITS sequencing, MassARRAY SNP mass spectrometry, Southern blotting, DNA conformation analysis and other molecular and biochemical approaches, it has been documented that *C. sinensis* comprises multiple intrinsic fungi from over 30 genera and at least 6 genotypes of *O. sinensis* [5–28]. An anamorph-teleomorph connection has been proposed for *Hirsutella sinensis* and *O. sinensis* [16,29]; however, no scientific studies to date have fully satisfied Koch's postulates by detailing the successful artificial induction of the *C. sinensis* sexual fruiting body and ascospores [23,29–36]. Based on the microcosmic results of the systematic and molecular systematic studies listed above, researchers have hypothesized that *C. sinensis* is an integrated micro-ecosystem with differential expressions of multiple intrinsic fungi (representing altered biomass in different compartments of natural *C. sinensis* during maturation, mainly caused by differential proliferations of intrinsic fungi plus many other biological processes) [18–28,33], and have

revealed a culture-dependent microbial community or mycobiota in wild *C. sinensis* with evidence for possible symbiotic interactions among the component fungi [23,37]. In addition to the microcosmic systematics techniques mentioned above, the random amplified polymorphic DNA (RAPD) technique has shown an advantage as a macrocosmic, fuzzy technique (not specific to particular sequences when viewing the study material as a whole) in molecular systematics for comparing overall polymorphism similarities and cluster relationships in different systems, in particular when studying natural *C. sinensis* with multiple known and unknown, culturable and non-culturable fungi. Using multiple random primers as required, RAPD studies have provided integrated analysis of the overall differences between comparison samples, contributing to a macrocosmic view in the molecular systematics along with the microcosmic evidence from the studies mentioned above. However, experiments using the macrocosmic RAPD technique have encountered several challenges and uncertainties, such as the maturational changes in the differential expression of the intrinsic *C. sinensis*-associated fungi, unbiased selection of multiple random primers, objective analysis of RAPD molecular marker polymorphisms and unbiased interpretation of the results [16,30,38–45]. One of the most common difficulties arises in the polymorphism similarity computation and cluster tree construction. The equation published by Nei and Li [46] is used to calculate similarity for “all or none” data in paired groups, particularly to analyze genetic variations through the use of restriction endonucleases, but it neglects the differences in DNA band densities, which contain a great deal of molecular information relating to the composition of the fungal community in

¹ The Latin name has been used indiscriminately for both the wild product and the teleomorph/holomorph of the *C. sinensis* fungus. The fungus was recently re-named *Ophiocordyceps sinensis* (Berk.) [5]. Because a consensus for a Latin name for the wild product has not been reached by taxonomists in mycology and TCM botany, we temporarily use the term “*O. sinensis*” to refer to the fungus and continue to use the name *C. sinensis* to refer to the wild product.

micro-ecosystems, such as *C. sinensis*, that comprise multiple intrinsic fungi. In this study, we formulated new similarity equations for the quantitative analysis of RAPD molecular marker polymorphisms. We also used density-weighted arithmetic methods for cluster tree constructions, compared to the unweighted algorithm that has been often used in prior *C. sinensis* studies.

2. Materials and Methods

2.1 Collection of *C. sinensis*

Fresh *C. sinensis* specimens were collected by local herb farmers from Kangding County in Sichuan Province (Qinghai-Tibetan Plateau area), China (Figure 1) and purchased locally. Premature (preM) *C. sinensis* specimens were collected in mid-May. These specimens featured a plump caterpillar body (sclerotium) and a short

stroma ranging from 1.0 to 2.0 cm long (1.5 cm on average). Maturing (M'ing) *C. sinensis* specimens featured a stroma approximately 3.0 to 5.0 cm long (4.0 cm on average). Mature (M) *C. sinensis* were collected in mid-June, having a less plump caterpillar body and a long stroma >5.0 cm (7.0 cm on average) with the formation of an expanded portion close to the tip, which is densely covered with ascocarp (AC). All freshly collected *C. sinensis* specimens were washed thoroughly on site in running water with gentle brushing, soaked in 0.1% mercuric chloride for 10 min for surface sterilization and washed 3 times with sterile water. The samples were immediately frozen in liquid nitrogen for transportation to the Lab in Beijing and storage prior to further processing [18].

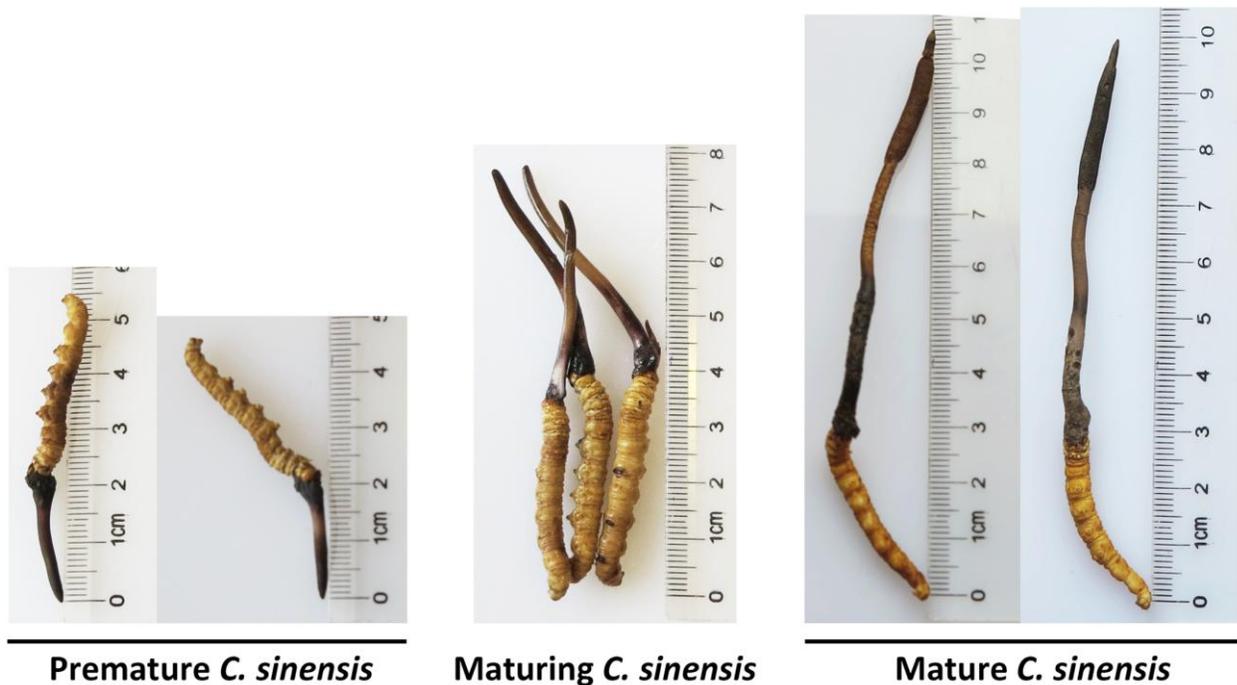


Figure 1. Fresh natural *C. sinensis* at the three stages of maturation with various heights of stromata.

2.2 Mycelia culture of *Hirsutella sinensis*

A strain of *Hirsutella sinensis* (CGMCCC Collection No. 7.7, China General Microbiological Culture Collection Center) was provided by Prof. Yinglan Guo (Institute of

Microbiology, China Academy of Sciences). *H. sinensis* was cultured on solid PDAP (PDA + 1% peptone) culture medium at 15-18 °C for 45 days.

2.3 Extraction of genomic DNA as the PCR templates

Ten cleaned *C. sinensis* specimens for each maturation stage were used for this study. The caterpillar body, stroma and ascocarp portion of each specimen and freeze-dried *H. sinensis* mycelia were ground into powder (individually) in liquid nitrogen according to maturation stage as caterpillar body, stroma and ascocarp portion samples. Genomic DNA was extracted from the above samples using the DNeasy Plant Mini Kit

(Qiagen) and quantified by measuring absorbance at UV₂₆₀ (Eppendorf BioSpectrometer®; Hamburg, Germany).

2.4 Random primers

For the purpose of comparing the density-unweighted and weighted algorithms in this study, 10 random primers were selected according to the literature (Table 1) [16,41]. They were also used in other studies [42-44]. The primers were synthesized by Invitrogen.

Table 1. Ten random primers selected according to the reference papers

Primer	Sequence	Reference	Also used in other studies	
S7	ggtgacgcag	[41]		[44]
S10 or P2	ctgctgggac	[41]	[43]	[44]
S23 or P3	agtcagccac	[16] [41]	[42]	[44]
S31 or P4	caatcgctg	[41]	[42]	[44]
S33 or P5	cagcaccac	[16] [41]	[42]	[44]
S61 or P6	ttcgagccag	[16] [41]	[42]	[44]
S69	ctcaccgtcc	[41]		
S80	acttcgccac	[41]		[44]
OPV14	agatcccgcc	[16]		
OPW06	aggcccgatg	[16]		

2.5 PCR amplification of DNA fragments and agarose gel electrophoresis

PCR experiments were performed as groups with each experiment using one of the random primers listed in Table 1 for the comparison samples. The genomic DNA (50 ng of each sample) isolated from the above described *C. sinensis* samples and *H. sinensis* mycelia as the templates were subjected to PCR amplification. PCR was run under the same conditions for the multiple comparison samples. The PCR reagent kits contained the following (Thermo, Foster City, USA): 5 µl 10X buffer, 4 µl 25 mmol MgCl₂, 1 µl 10 mmol dNTP, 0.5 µl 100 µM primer, 0.5 µl 5 U/µl Taq, 50 ng DNA, and adding H₂O to 50 µl (final volume). The PCR protocol was as follows: 95 °C for 5 min; 40 cycles of 94 °C for 1 min, 36 °C for 1 min, and 72 °C for 2 min; then 72 °C for 2 min; and a hold at 4 °C. For comparison purposes, all PCR products (10 µl of each sample) from a single-

primer PCR experiment were applied to the lanes of agarose gel electrophoresis (1.5% agarose) performed at 120 V for approximately 2 hours. The agarose gels were stained with ethidium bromide and de-stained with water.

2.6 RAPD polymorphism similarity calculation

The agarose gels were photographed and scanned using Quantity One software (Bio-Rad) for comparisons of the DNA bands in the different samples. From the images of the scanned RAPD gels, the software digitalized the density and *R_f* (Relative front) of all amplicons and estimated their molecular sizes relative to the molecular weight standard. Qualitative (presence or absence) and quantitative (density) information of all DNA bands in the gel images were used in the polymorphism similarity comparisons and cluster construction. Using the density-

unweighted Nei-Li equation [46] to calculate polymorphism similarity (S) between any pairs of the samples:

Nei-Li [46] density-unweighted equation:

$$S = (2 N_{AB}) / (N_A + N_B) \quad (1)$$

where N_{AB} refers to matched DNA bands in lanes A and B amplified through PCR and N_A and N_B refer to the total number of DNA bands in lanes A and B.

In the contrast to the similarities calculated using the unweighted Nei-Li [46] equation (1), new ZUNIX equations (*cf.* the Results section) were formulated for similarity computation with amplicon-density weighting (Beijing Bioland Technology, 2013; <http://www.ebioland.com/ZUNIX.htm>), considering (i) mismatched DNA bands and (ii) matched but dissimilar DNA bands. A prior normalization was applied when computing the similarities of the molecular markers integrated from multiple gels, and the total density of a molecular weight standard (DNA ladder) was used as the normalization reference.

2.7 Density-weighted cluster analysis for polymorphic RAPD DNA

The normalized polymorphic molecular data were used for cluster construction. For mismatched DNA pairs, a missing band at the given location (molecular size) was assigned a score of 0. The digital density data of all matched and unmatched DNA bands in the lanes of comparison were ranked and arbitrarily assigned a score of 1 through 9 (the density-weighted, semi-quantitative algorithm) in accordance with the rank of their densities in 2 or more sample lanes of comparison, or assigned a score of 1 when the density-unweighted algorithm was used. The neighbor-joining distance method provided by the PAUP 4.0B software (Swofford, 2002) was used to compare the density-unweighted and weighted semi-quantitative algorithms for cluster construction (bootstrap = 1000). In addition to the semi-quantitative algorithm provided by PAUP 4.0B, full-quantitative cluster analysis was also performed using parametric hierarchical clustering algorithms; these include Pearson correlation

(uncentered), absolute correlation (uncentered), Euclidean distance, Squared Euclidean Distance, City-block distance provided by Cluster 3.0 (Stanford University, CA and Human Genome Center, University of Tokyo), JMP9 (The SAS Institute, Cary, NC) and SPSS 10.1 (SPSS Inc., Chicago, IL), including different distance arithmetic methods such as average linkage between groups, average linkage within-group, Ward and Fast Ward linkage, nearest neighbor (single linkage), furthest neighbor (complete linkage), and centroid linkage distance.

3. Results

3.1 Extreme cases of RAPD gels with the use of S31 and S69 primers

Figure 2 shows two RAPD gel images of the DNA bands amplified using Primers S31 and S69. Calculated using the Nei-Li [46] equation (1), a similarity of 42% was found between the ascocarp portion and the *H. sinensis* samples using the S31 primer (the left panel); and a similarity of 50% was found between the mature caterpillar body and premature stroma samples using the S69 primer (the right panel).

3.2 New equations for computing the similarity of RAPD polymorphisms

The Nei-Li [46] equation (1) is used for similarity computations and is ideal for "all or none" data under 2 prerequisites: (a) all matched DNA pairs in the electrophoretic lanes of comparison are identical with essentially the same densities, and (b) all amplicons are well separated from the adjacent DNA moieties with similar molecular weights by electrophoresis. However, in reality as shown in Figure 2, many matches have similar or dissimilar densities for the DNA bands of the pairs of electrophoretic lanes, and some amplicons are not well separated from the adjacent DNA moieties. On the left panel of Figure 2, the estimated 1707-bp and 557-bp bands in the ascocarp (AC) lane and the 2078-bp band in the *H. sinensis* (Hs) lane should be much heavier weighted in the similarity

calculation. In addition, on the right panel of Figure 2, the matched, high density 1473-bp bands in both lanes should also be given substantially heavier weight in the similarity calculation, while those mismatches with much less density should not be weighted equally as the 1473-bp bands in similarity calculations or in cluster construction. Due to non-satisfaction of the prerequisites, the Nei-Li [46] equation (1) is not suitable for similarity computations for the RAPD studies of *C. sinensis* and results in a loss of the wealth of molecular information provided by the density of the DNA amplicons during RAPD polymorphism analysis under a primitive understanding of the reality. New arithmetic methods for similarity computation are required to quantify the molecular information provided by the density of the DNA amplicons, which are lost or partially lost during RAPD polymorphism analysis using the Nei-Li equation (1) [46]. Accordingly, the ZUNIX equation (2) was formulated to compare the polymorphisms of 2 electrophoretic lanes: $d_{ik} \geq 0, i=1,2, k=1,2, \dots, m$. We defined the measure of similarity as follows:

$$S = \frac{\sum_{k=1}^m [2 \text{Min}\{d_{1k}, d_{2k}\}]}{\sum_{k=1}^m (d_{1k} + d_{2k})} \quad (2)$$

where the similarity of the 2 densities d_{1k} and d_{2k} is the common portion of their values. The ZUNIX equation (2) defines similarity as the total density of all common parts present in the samples of comparison divided by the total density of all bands across the samples. The ZUNIX equation (2) is mathematically general with no specific prerequisites and governs all conditions, in contrast to the density-unweighted Nei-Li [46] equation (1) that narrows the specific

cases under the strict prerequisites (a) and (b) described above.

Using the density-weighted ZUNIX equation (2), the similarity of 42% between the ascocarp portion and *H. sinensis* samples calculated using Nei-Li [46] equation (1) was accurately reduced to 16.8% (-60%) using the S31 primer (*cf.* the left panel of Figure 2); the similarity of 50% between the mature caterpillar body and premature stroma samples was accurately increased to 89.5% (+79%) using the S69 primer (*cf.* the right panel of Figure 2). These results indicate that the density-unweighted Nei-Li [46] equation (1) is not suitable for molecular systematic RAPD studies for natural *C. sinensis* and that the density-weighted ZUNIX equation (2) accurately captured all of the molecular information buried in the amplicon DNA bands (both the density and the migration speed in the agarose gel electrophoresis) in the RAPD gel images.

The second ZUNIX equation (3) is suitable for comparing the DNA amplicons in 3 electrophoretic lanes, where $d_{ik} \geq 0, i=1,2,3, k=1,2, \dots, m$, and is shown below.

$$S = \frac{\sum_{k=1}^m [3 \text{Min}\{d_{1k}, d_{2k}, d_{3k}\}]}{\sum_{l=1}^m (d_{1l} + d_{2l} + d_{3l})} \quad (3)$$

Extending further, the third ZUNIX equation (4) is suitable for comparisons of DNA amplicons in more than 3 lanes, where $d_{ik} \geq 0, i=1,2, \dots, n, k=1,2, \dots, m$, and is described as:

$$S = \frac{\sum_{k=1}^m [n \text{Min}\{d_{1k}, d_{2k}, \dots, d_{nk}\}]}{\sum_{r=1}^n \sum_{s=1}^m d_{rs}} \quad (4)$$

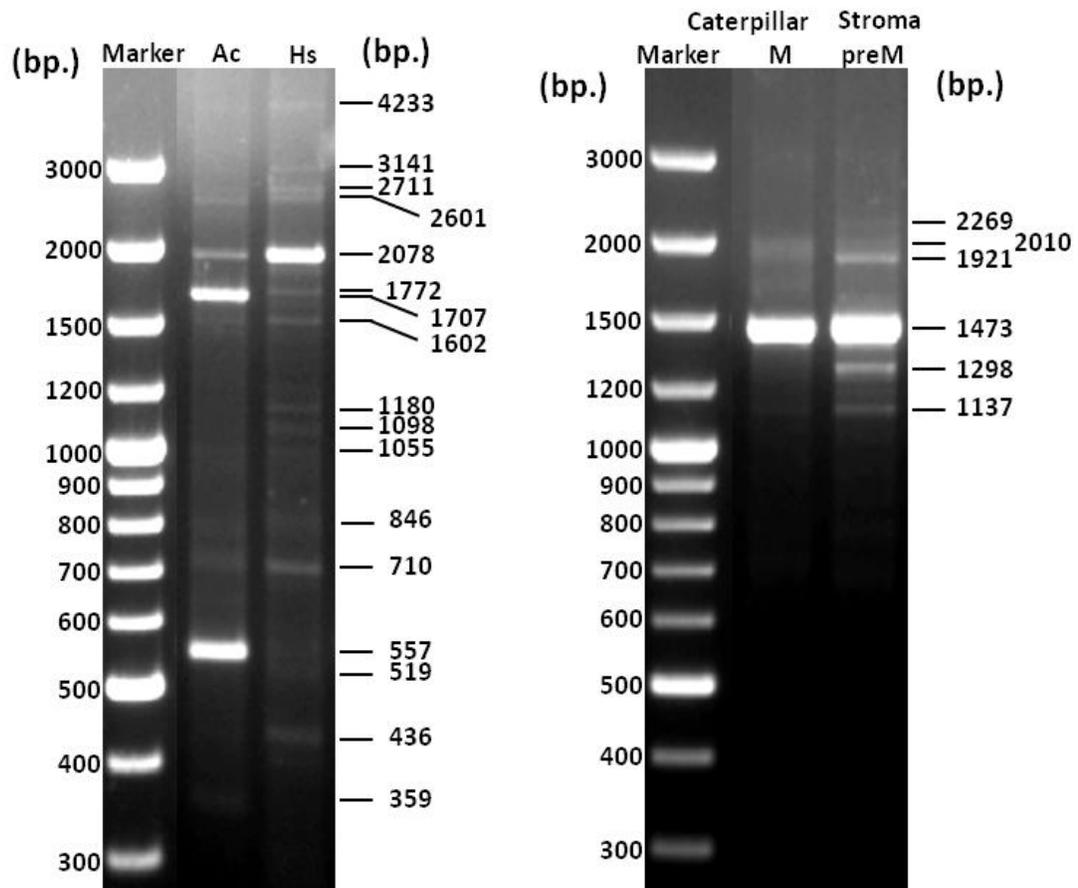


Figure 2. Agarose gel electrophoresis of S31 and S69 primers RAPD PCR amplicons. DNA amplicons were amplified by PCR using random primers S31 (left panel) or S69 (right panel). The genomic DNA used as the template for PCR was isolated from the ascocarp portion of wild *C. sinensis* and *H. sinensis* mycelia for the experiment shown on the left panel, and from the mature caterpillar body and premature stroma of wild *C. sinensis* for the experiment shown on the right panel. AC refers to the ascocarp portion; Hs refers to the *H. sinensis* mycelia; M and preM refer to mature and premature *C. sinensis*, respectively.

3.3 RAPD polymorphisms with the use of 10 random primers

RAPD experiments were performed using the 10 random primers listed in the Materials and Methods to obtain 10 RAPD image profiles. These data were used to compare the performances of the density-weighted and unweighted algorithms for similarity computation and cluster construction. Figures 3 and 4 show the images of the DNA bands amplified using the S7 and S31 primers, respectively, as the examples

of the 10 gel images. Both figures demonstrate largely diversified RAPD molecular marker polymorphisms.

3.4 Similarities between ascocarp portion and caterpillar/stroma samples computed using the Nei-Li equation and the new equation

In contrast to the Nei-Li [46] equation (1) that is suitable for “all or none” data analysis, the new ZUNIX equation (2) quantitatively integrates

the density differences of the amplicons in the pair of comparison lanes into the similarity computation; thus, it retains as much the molecular information from the RAPD gel images as possible, which is extremely important when macrocosmically comparing samples with differential expressions of multiple intrinsic

fungi. Table 2 shows the similarities of the RAPD polymorphisms as measured using the S7 or S31 primer between the ascocarp portion and other *C. sinensis* samples at the 3 maturation stages with the Nei-Li [46] equation (1) and the ZUNIX equation (2) for 2-lane data (cf. Figures 3&4).

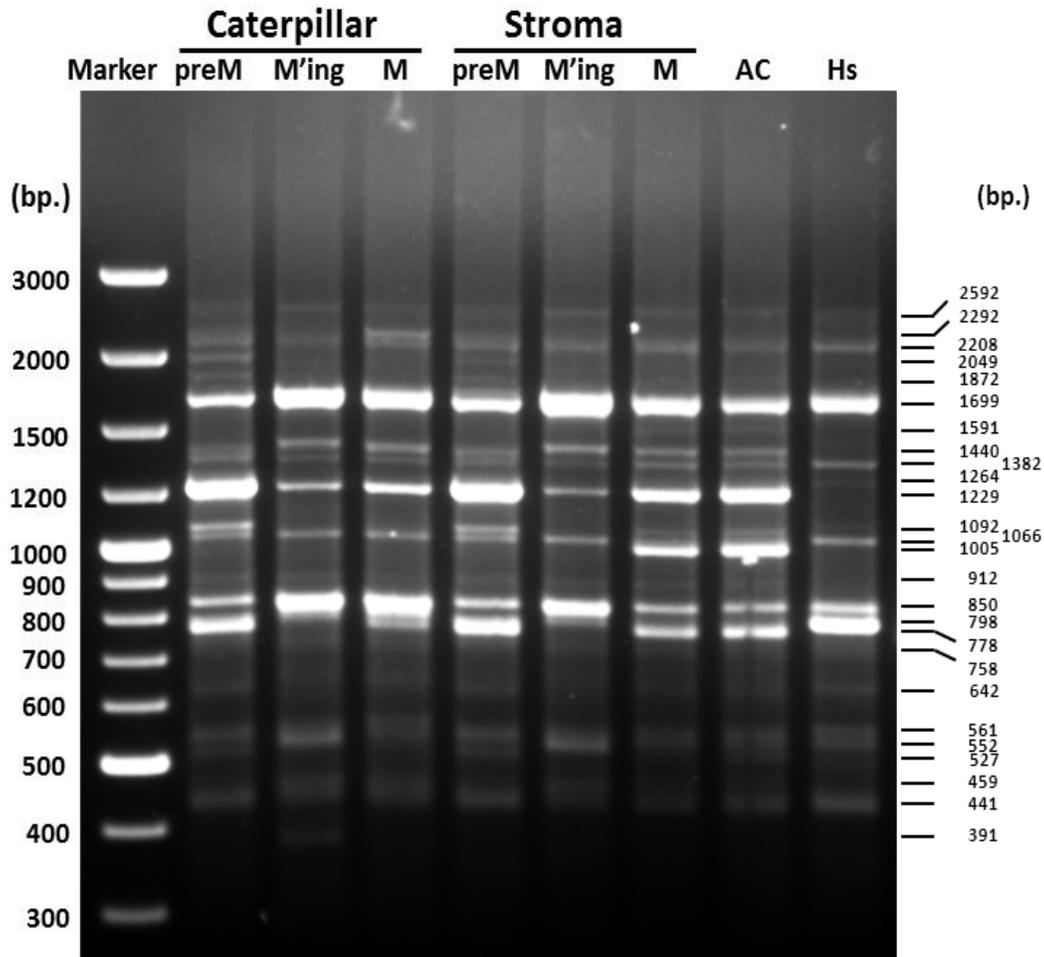


Figure 3. Agarose gel electrophoresis of S7 primer RAPD PCR amplicons. DNA amplicons were amplified by PCR using primer S7. The genomic DNA used as the template for PCR was isolated from the caterpillar bodies, stromata or ascocarp portion of wild *C. sinensis* at 3 maturation stages or from *H. sinensis* mycelia. preM refers to premature *C. sinensis*; M'ing refers to maturing *C. sinensis*; M refers to mature *C. sinensis*; AC refers to the ascocarp portion of mature *C. sinensis*; and Hs refers to the *H. sinensis* mycelia.

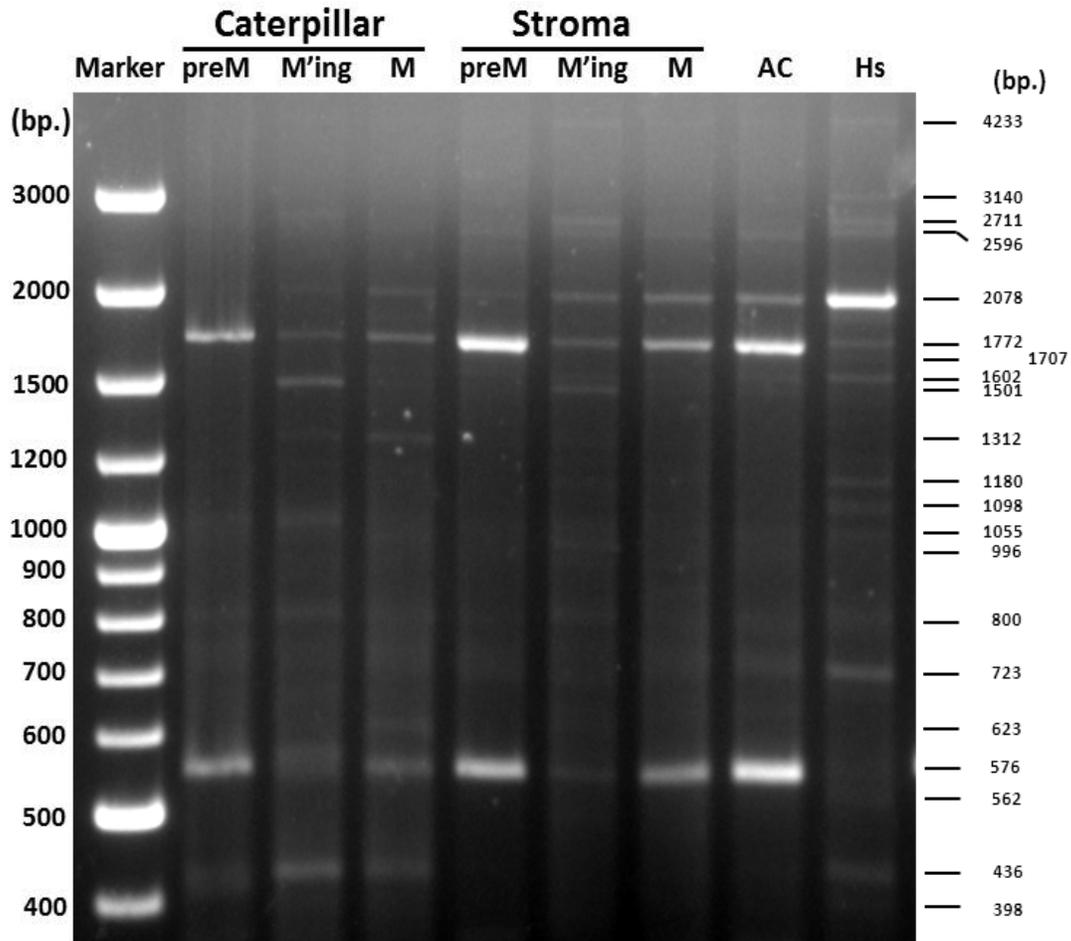


Figure 4. Agarose gel electrophoresis of S31 primer RAPD PCR amplicons. DNA amplicons were amplified by PCR using primer S31. The genomic DNA used as the template for PCR was isolated from the caterpillar bodies, stromata or ascocarp portion of wild *C. sinensis* at 3 maturation stages or from *H. sinensis* mycelia. preM, M'ing and M refer to premature, maturing and mature *C. sinensis*; AC refers to the ascocarp portion of mature *C. sinensis*; and Hs refers to the *H. sinensis* mycelia.

Table 2. Similarities of RAPD molecular marker polymorphisms between the ascocarp portion and other *C. sinensis* samples computed using the Nei-Li [46] equation (1) and the new ZUNIX equation (2).

Primer	Similarity computed with the	Polymorphism similarity (%) between the ascocarp and other <i>C. sinensis</i> samples					
		Caterpillar Body			Stroma		
		Premature	Maturing	Mature	Premature	Maturing	Mature
S7	Nei-Li equation	85.7%	56.0%	76.9%	88.9%	83.3.9%	96.0%
	ZUNIX equation	62.9%	44.3%	45.3%	56.7%	36.9%	89.8%
S31	Nei-Li equation	58.8%	31.6%	47.1%	85.7%	55.6%	76.9%
	ZUNIX equation	66.6%	3.3%	25.8%	70.5%	13.2%	55.9%

As shown in Table 2, the similarities of the primer S7 RAPD polymorphisms computed using the ZUNIX equation (2) were 36.9% to 89.8% between the ascocarp portion and other *C. sinensis* samples. The similarities of the primer S31 RAPD polymorphisms computed with the ZUNIX equation (2) were 3.3% to 70.5% between the ascocarp portion and other *C. sinensis* samples. In general, the mass majority of these similarities are smaller than those computed using the Nei-Li [46] equation (1). One exception was observed when comparing the primer S31 amplicons from the ascocarp portion with those from the premature caterpillar body samples; in that case, the similarity computed with the ZUNIX equation (2) was larger than that computed with the Nei-Li [46] equation (1). This result is obtained because the densities of the 2 matched DNA amplicons, estimated 1772 bp and 576 bp (*cf.* Figure 4), are weighted much more heavily than the data for the unmatched DNA bands. Clearly from the results, the ZUNIX equation (2) is capable of accurately capturing quantitative differences in DNA band density,

whereas the Nei-Li [46] equation (1) is not. Table 2 also shows large differences in RAPD polymorphisms between the ascocarp portion and other compartments of natural *C. sinensis* when using different random primers.

3.5 Similarities between *H. sinensis* and *C. sinensis* samples computed with use of the new ZUNIX equation

The similarities of the RAPD polymorphisms between the *H. sinensis* and *C. sinensis* samples computed with the Nei-Li and the ZUNIX equations are shown in Table 3. The similarity of the S7 RAPD polymorphisms computed with the ZUNIX equation (2) was 41.2% between *H. sinensis* and the ascocarp portion sample, and 46.3% to 71.2% between *H. sinensis* and other *C. sinensis* samples (Table 3). The similarity of the S31 RAPD molecular marker polymorphisms computed with the ZUNIX equation (2) was 16.8% between *H. sinensis* and the ascocarp portion sample, and 0.6% to 18.9% between *H. sinensis* and other *C. sinensis* samples.

Table 3. RAPD molecular marker polymorphism similarities between *H. sinensis* and *C. sinensis* samples computed with the Nei-Li [46] equation (1) and the new ZUNIX equation (2).

Primer	Similarity computed with the	RAPD polymorphism similarity (%) between <i>C. sinensis</i> samples and <i>H. sinensis</i>						
		Caterpillar Body			Stroma			Ascocarp portion
		Premature	Maturing	Mature	Premature	Maturing	Mature	
S7	Nei-Li equation	69.2%	60.9.3%	83.3%	88.0%	63.6%	87.0%	83.3%
	ZUNIX equation	50.3%	53.7%	48.7%	71.2%	50.4%	46.3%	41.2%
S31	Nei-Li equation	21.1%	47.6%	42.1%	50.0%	60.0%	40.0%	33.3%
	ZUNIX equation	4.1%	1.5%	1.3%	0.6%	18.9%	16.5%	16.8%

Figures 3&4 also showed the large difference in the total density of electrophoretic lanes between the S7 and S31 gels, indicating the necessity of across-gel normalization prior to integral density-weighted analyses. All the DNA polymorphism information from the 10 RAPD gels was normalized using the total densities for

the molecular weight standard (DNA ladder) as the reference and then integrated. Using the ZUNIX equation (2), Table 4 lists the integral similarities of the RAPD polymorphisms between the ascocarp portion and other *C. sinensis* samples, which ranged from 54% to 88%, indicating the diverse distribution of intrinsic

fungi in the *C. sinensis* compartments and the changes in their expression during *C. sinensis* maturation. We also observed much lower integral similarities of RAPD polymorphisms between the *H. sinensis* and *C. sinensis* samples,

which ranged from 46% to 66%, specifically 65% between the *H. sinensis* and ascocarp portion sample, indicating significant dissimilarity between *H. sinensis* and *C. sinensis*.

Table 4. Integral similarities of RAPD polymorphisms computed all 10 RAPD gels using the new ZUNIX equation (2)

Integral similarity	Caterpillar Body			Stroma			Ascocarp portion
	Premature	Maturing	Mature	Premature	Maturing	Mature	
Ascocarp portion vs. other <i>C. sinensis</i> samples	65.0%	54.4%	66.2%	73.9%	56.9%	87.7%	
<i>H. sinensis</i> vs. <i>C. sinensis</i> samples	51.9%	45.8%	55.6%	65.2%	47.2%	65.5%	65.1%

3.6 Integral polymorphism similarities among *C. sinensis* samples during maturation

By further expanding the capability of the similarity computation, the new ZUNIX equation (3) was used to compute the similarities for 3 sample lanes. We found differences in the RAPD polymorphisms of the caterpillar bodies at the 3 maturation stages (*cf.* Figures 3&4 as the examples), with a 52.4% integral similarity of the RAPD polymorphisms computed from the digital data of all 10 RAPD gel images. When the new ZUNIX equation (4) was used to compute the similarity of polymorphisms for more than 3 sample lanes, we found changes in the RAPD polymorphisms in the stromata of the 3 stages of maturation and ascocarp portion samples (*cf.* Figures 3&4 as the examples), with a 48.6% integral RAPD polymorphism similarity. The integral similarity of RAPD polymorphisms was 40% for all the *C. sinensis* samples (the caterpillar body, stroma and ascocarp portion) during maturation. This low value suggests diversified fungal background with significant changes in the differential fungal expression of the multiple intrinsic fungi in these compartments of *C. sinensis* during maturation.

3.7 Comparison of cluster trees constructed using the density-unweighted and weighted algorithms

An unweighted neighbor-joining method was used to construct the cluster tree (Figure 5) using the data from the primer S31 RAPD gel shown in Figure 4. The unweighted nature of this algorithm determined that the different densities of the DNA amplicons were not weighted in the algorithm, leading to a loss of important molecular information from the cluster analysis as shown in Figure 5. We then compared the cluster tree patterns using the density-unweighted and semi-quantitative density-weighted cluster algorithms provided in the same software, PAUP 4.0B. Using a semi-quantitative scoring of 1 to 9 for the density of the DNA amplicons, the density-weighted algorithm provided a more accurate way to capture the quantitative differences in the construction of a cluster tree for the S31 RAPD gel and obtained a completely different pattern of cluster tree (Figure 6), compared to Figure 5 constructed using the unweighted algorithm. This comparison demonstrated that the basic clade grouping was altered when switching the unweighted to weighted algorithms provided by the same software.

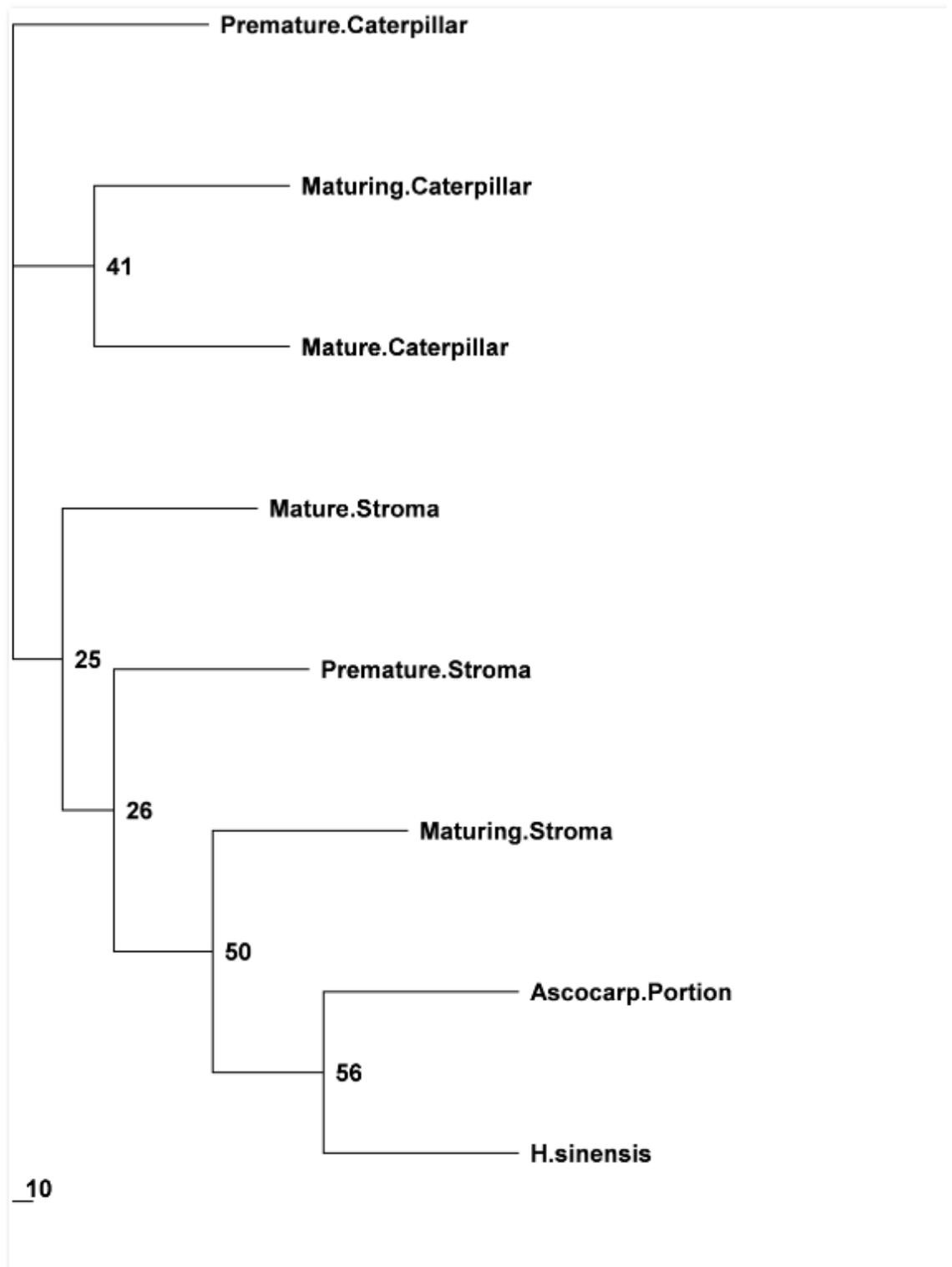


Figure 5. Cluster tree for the S31 RAPD gel constructed using the unweighted neighbor-joining algorithm. A cluster tree for the S31 RAPD gel shown in Figure 4 was constructed using the unweighted neighbor-joining algorithm method. Each DNA amplicon is considered as one character and assigned a score of 1; the missing DNA band at the same molecular weight is assigned a score of 0. (Neighbor-Joining distance method; bootstrap = 1000).

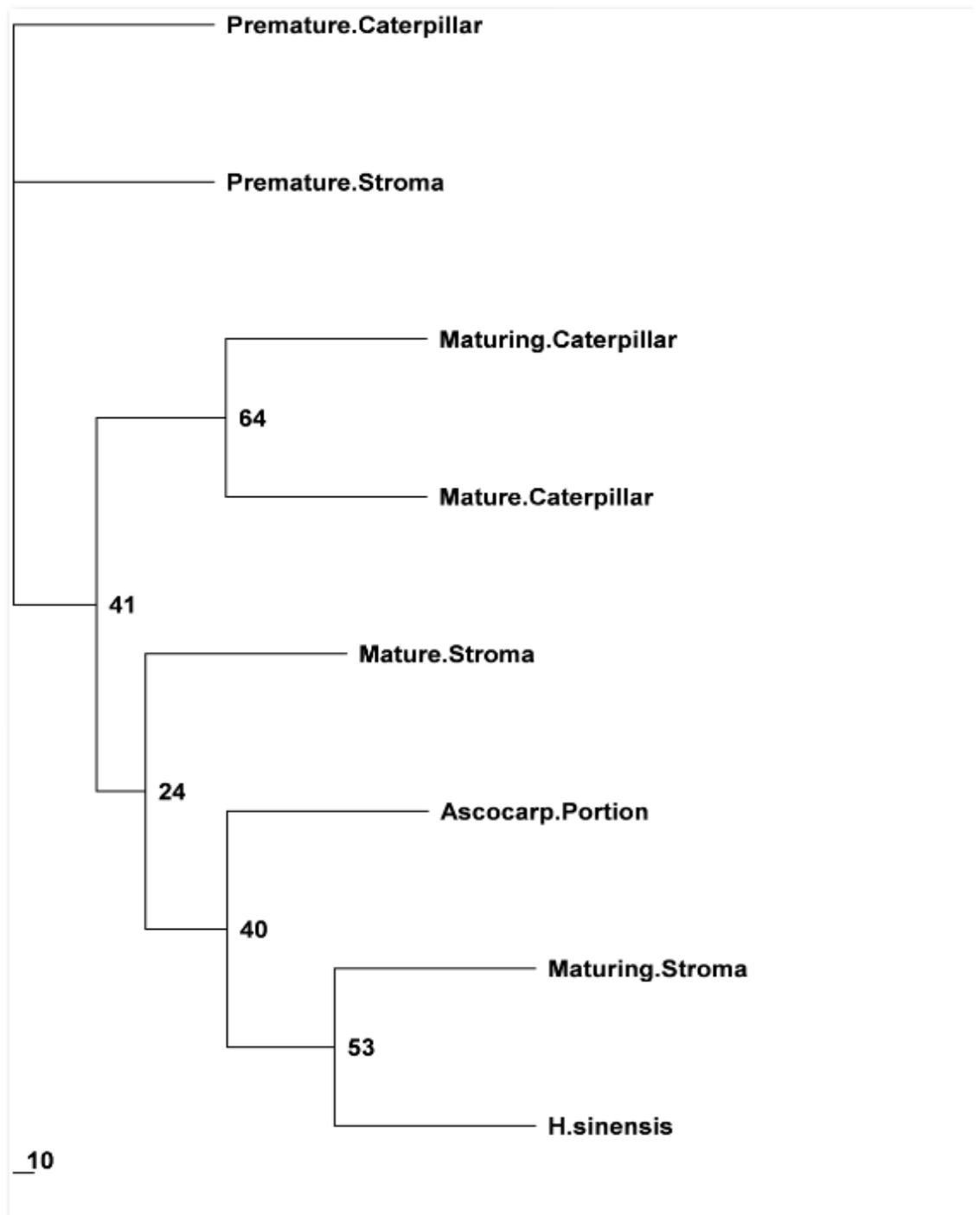


Figure 6. Cluster tree for the S31 RAPD gel constructed using the weighted algorithm. A cluster tree for the S31 RAPD gel shown in Figure 4 was constructed using the density-weighted algorithm method. Each DNA amplicon is considered as one character and assigned a score of 1-9 based on its density rank among the densities of all the DNA amplicons being compared; the missing DNA band at the same molecular weight is assigned a score of 0. (Neighbor-Joining distance method; bootstrap = 1000).

3.8 Integrated cluster trees constructed using the density-weighted algorithms

After normalization to the density of the total molecular standard (DNA ladder), an

integral cluster tree was constructed based on the RAPD molecular information from all 10 primer experiments with the density-weighted algorithm. As shown in Figure 7, when *H. sinensis* was not

included in the cluster analysis, 4 major RAPD clades were formed: Clade A, the premature stroma and caterpillar body formed an outside clade; Clade B, the mature caterpillar body; Clade C, the maturing caterpillar body/stroma; and Clade D, mature stroma and ascocarp portion. Clades B&C formed a small cluster,

which formed a larger cluster with Clade D. *H. sinensis* was added to the cluster analysis (Figure 8) and is situated closer to the small cluster comprised with Clades B&C, although the bootstrap value was only 50%. It is apparent that *H. sinensis* is further from the Clades A&D, and in particular is far from the ascocarp portion.

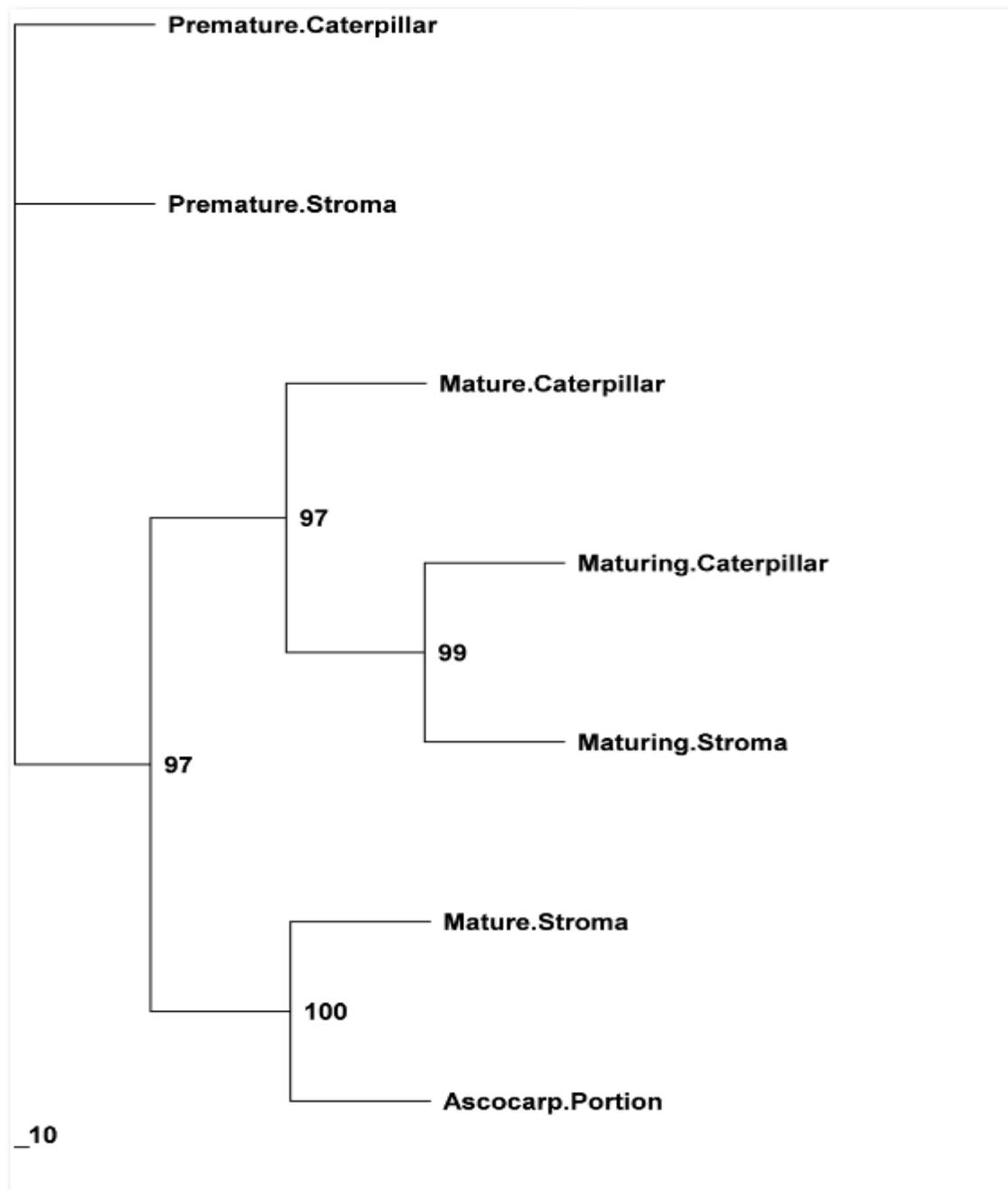


Figure 7. Integral cluster tree for all 10 RAPD gels constructed not including *H. sinensis*. An integral cluster tree for all 10 primer RAPD gels was constructed using the semi-quantitative density-weighted algorithm (PAUP 4.0B). Every DNA amplicon from every RAPD gel was considered as one character and assigned a score of 1-9 based on its density rank among the densities of all the DNA amplicons in all RAPD gels being compared; the missing DNA band at the same molecular weight was assigned a score of 0. (Neighbor-Joining distance method; bootstrap = 1000).

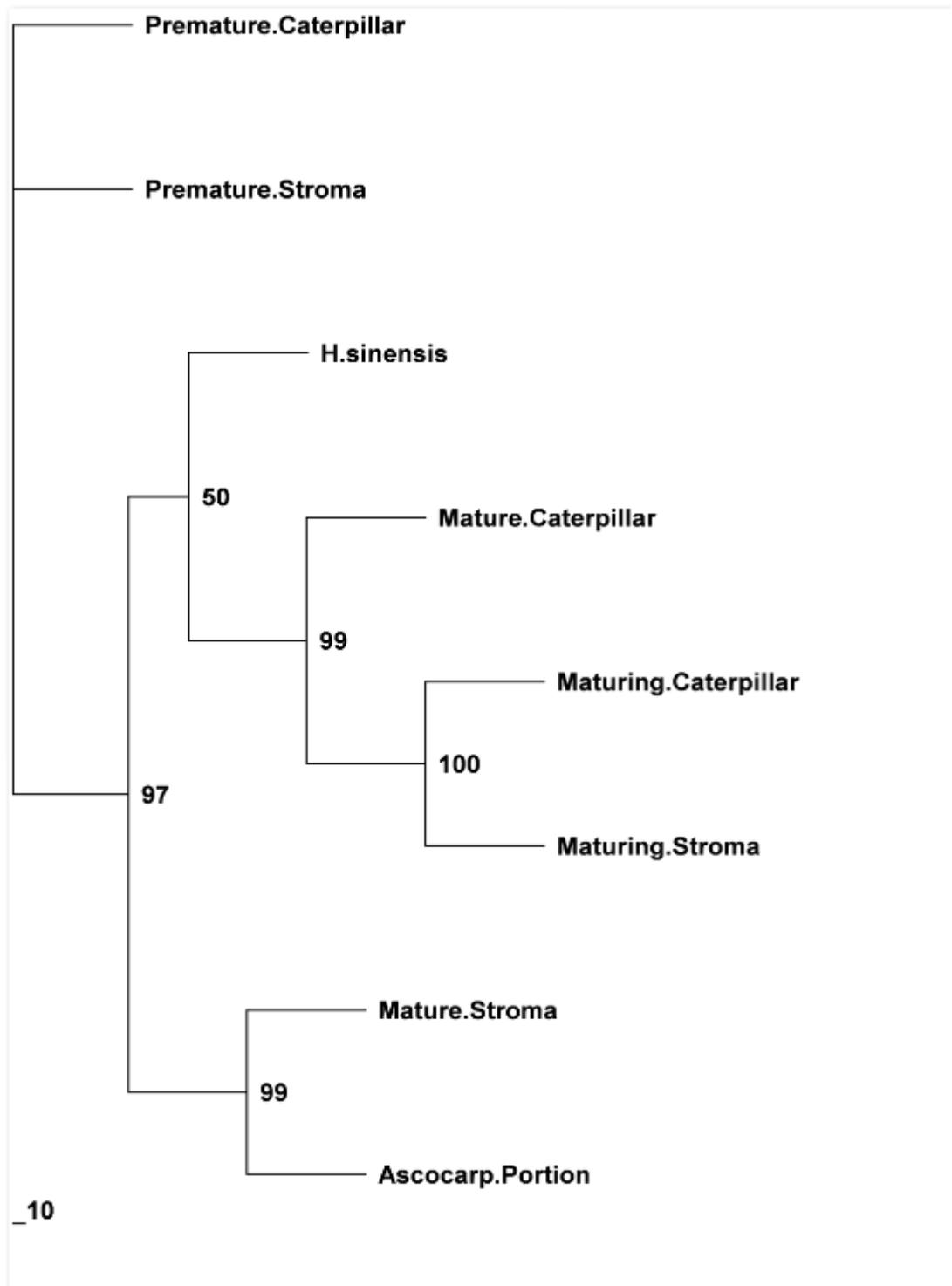


Figure 8. Integral cluster tree for all 10 RAPD gels constructed including *H. sinensis*. An integral cluster tree for all 10 primer RAPD gels was constructed using the semi-quantitative density-weighted algorithm (PAUP 4.0B). Every DNA amplicon from every RAPD gel was considered as one character and assigned a score of 1-9 based on its density rank among the densities of all the DNA amplicons in all RAPD gels being compared; the missing DNA band at the same molecular weight was assigned a score of 0. (Neighbor-Joining distance method; bootstrap = 1000).

3.9 Integrated cluster trees constructed using the fully-quantitative density-weighted algorithms

The above use of PAUP 4.0B software allowed us to directly compare the cluster trees constructed using the density-unweighted and weighted algorithms, demonstrating that the basic grouping (clade formation) for the samples based on the RAPD polymorphisms was inaccurately constructed in the density-unweighted cluster tree (comparing Figures 5 & 6). However, this software provides only limited ability in handling full-quantitative density data, and we can only use a scoring system for semi-quantitative clustering analyses. Thus the clustering analysis was further extended using other software with

full-quantitative capacity for cluster constructions, such as Cluster 3.0, JMP9 and SPSS 10.1. Figure 9 shows a cluster tree as an example, using the SPSS centroid linkage distance method of hierarchical cluster analysis. Comparing the cluster trees constructed using the fully quantitative density-weighted algorithms provided by the different software (see Method section for details), the basic patterns of clade formation are similar in the trees shown in Figures 8 and 9 constructed by the PAUP 4.0B semi-quantitative neighbor-joining distance algorithm and the SPSS 10.1 fully quantitative centroid distance algorithm. Four basic clades consistently appear in the cluster trees for the *C. sinensis* samples, as described above.

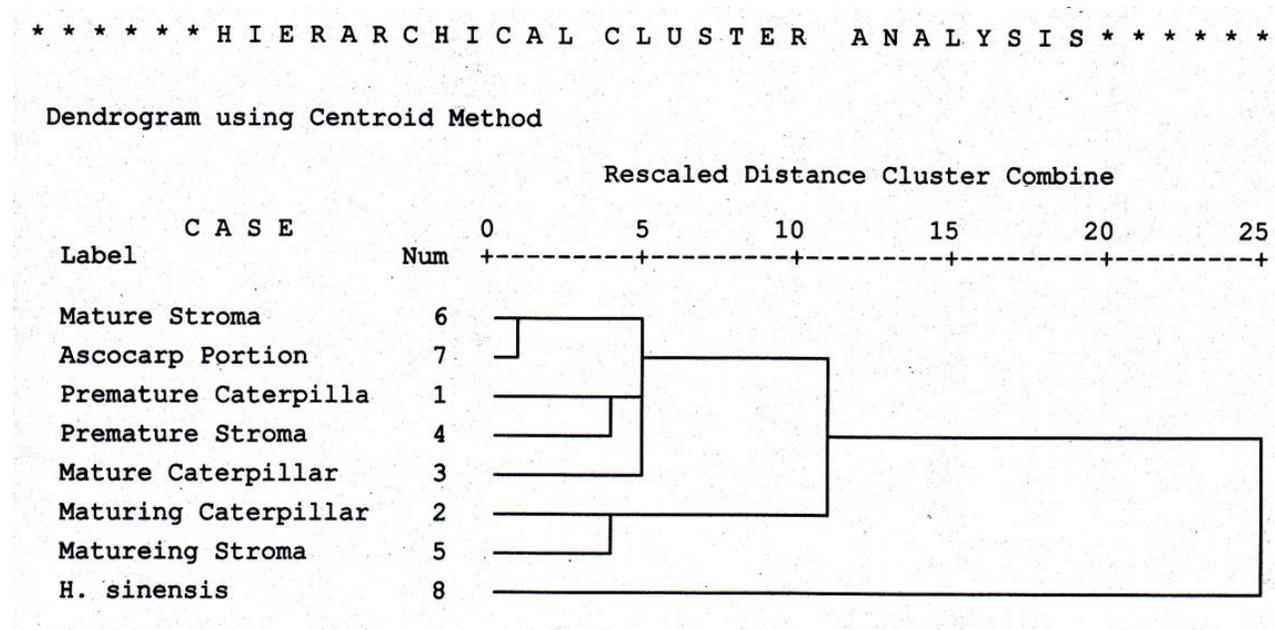


Figure 9. Integral cluster tree for all 10 RAPD gels constructed with a fully quantitative algorithm. An integral cluster tree for all 10 primer RAPD gels was constructed using the fully quantitative density-weighted, parametric algorithm provided by SPSS 10.1 software. Normalized density data for all DNA amplicons in all RAPD gels were entered into the cluster construction (Centroid Linkage distance method).

Several differences in the cluster tree patterns were noted when comparing the cluster trees constructed by the various quantitative algorithms, including the semi-quantitative

method. (i) With the semi-quantitative algorithm (*cf.* Figures 7&8), Clade B (mature caterpillar body) is closer to Clade C (maturing caterpillar body and stroma), forming a small cluster, but is

distant from Clades A (premature caterpillar body and stroma) and D (ascocarp portion and mature stroma). However, with the various fully quantitative algorithms (cf. Figure 9 for the example), Clade B is closer to Clades A and D, forming a small cluster. This small cluster is in a short distance from Clade C. (ii) Using the semi-quantitative algorithm (cf. Figure 8) and some (Word and fast Word Linkage, Pearson Correlation of Average Linkage (Within Group)) of the fully quantitative algorithms, *H. sinensis* was placed inside of the overall *C. sinensis* cluster and closer to the small *C. sinensis* cluster comprised of Clades B and C. However, using other fully quantitative algorithms (cf. Figure 9), *H. sinensis* is in a long distance from the overall *C. sinensis* cluster comprised of all *C. sinensis* clades. (iii) Different patterns of cluster trees can be seen constructed with the same algorithm; one example is the Word Linkage distance method provided by different software (JMP vs. SPSS). However, the similar cluster patterns can be seen constructed with the semi-quantitative Neighbor-Joining distance algorithm provided by PAUP 4.0B and the fully quantitative City-Block distance algorithm provided by Cluster 3.0 (data not shown). As demonstrated through the above comparisons, it is notable that the most important aspect for clustering analysis on *C. sinensis* RAPD polymorphism data is the basic clade formation (grouping), which appeared to be largely inaccurate when using density-unweighted algorithm based on comparisons between Figures 5&6. Without changing the basic clade formation, all the quantitative density-weighted algorithms tested, including the semi- and fully quantitation, may construct cluster trees with somewhat altered between-clade distances due to the various distance arithmetic methods used.

4. Discussion

C. sinensis grows only in areas of high elevation on Qinghai-Tibetan Plateau, and has a

complex life cycle [1–2]. The season-/temperature-dependent features of *C. sinensis* maturation have long been recognized by traditional Chinese medicine, which reports differences in the therapeutic potency of TCM *C. sinensis* at different maturation stages. We have previously reported dynamic changes in the profiles and activity of the fungal species and in the component chemicals and proteins/peptides throughout *C. sinensis* maturation [18,24–28]. Many studies have reported that *C. sinensis* contains fungi from over 30 genera and at least 6 genotypes of *O. sinensis* and that the most abundant culturable fungi are *Pseudogymnoascus roseus* in both the sclerotia and cortices of *C. sinensis* and *Penicillium chrysogenum* in the stromata [6–30]. Because of the complex fungal genetic background of natural *C. sinensis*, global genome-wide sequencing failed due to problems associated with sequence assembly (reported by Prof. Chengshu Wang in the 10th Symposium of Mycology Societies cross Taiwan Strait, Wuhan China, July 15-17, 2011 and the 7th International Medicinal Mycology Conference, Beijing China, August 26-29, 2013); however, global genome-wide sequencing could be successfully performed on a purified fungus strain of *H. sinensis* (*Ophiocordyceps sinensis* was used as the fungus name in the paper) [48]. We certainly hope that the complete genome information will be available in the near future for more component fungi of natural *C. sinensis*. The fungal genetic background is even more complex when non-culturable fungal species are taken into consideration [21,24-28,47]. *C. sinensis* is therefore hypothesized as an “integrated micro-ecosystem” [33]. In addition to the microcosmic sequencing studies for molecular systematics, it is necessary to use macrocosmic molecular techniques, such as RAPD, to profile the global fungal molecular markers of all the intrinsic fungi in *C. sinensis*, although there are some limitations for the integral, fuzzy technology. The unweighted algorithms, Nei-Li [46] equation (1) for similarity computation and unweighted algorithms for cluster construction, have been used in previous *C. sinensis* RAPD studies [16,38–45]. However, these unweighted arithmetic methods are only truly suitable for the

analysis of “all-or-none” qualitative data when the data meet the 2 prerequisites: (a) all matched DNA pairs in the electrophoretic lanes of comparison are identical with similar densities, and (b) all amplicons are well separated from the adjacent DNA moieties with similar molecular weights and conformations by electrophoresis. In contrast to the narrow, specific applicability set by the prerequisites, the density-weighted, quantitative arithmetic methods formulated in this study are more general mathematically with no restricted prerequisites and capture all the detailed molecular information related to the dynamic changes in the fungal expression and relative abundance of the various intrinsic fungi during *C. sinensis* maturation, as shown in Tables 2 and 3 and in previous publications [18,21,24-28]. Most of the similarities computed with the new ZUNIX equation (2) are fewer than those calculated using the Nei-Li [46] equation (1), because the new algorithm considered not only the qualitative data (i.e., the presence or absence of matching bands), but also the density differences of the amplicons, particularly the most abundant DNA bands (*cf.* Tables 2&3). However, when the S31 amplicons in the ascocarp portion and the premature caterpillar body samples were compared, the similarity computed using the ZUNIX equation (2) was greater than that calculated with the Nei-Li [46] equation (1), because the densities of 2 matched DNA bands, 1772 bp and 576 bp as estimated relative to the molecular weight standards, are weighted much more heavily than the unmatched DNA bands (*cf.* Figure 4). The clade formation in the S31 cluster tree constructed using the unweighted method that overlooked the density differences of the amplicons (*cf.* Figure 5) appears very different from that constructed with the density-weighted neighbor-joining method considering the weighting of all the amplicon densities (*cf.* Figure 6). These results demonstrated that from two angles of global profile comparisons that inter-support each other, the density-weighted algorithms for the similarity computation and for cluster construction provide an apparent advantage in accuracy for polymorphism data analysis in *C. sinensis* containing multiple intrinsic fungi.

Jiang and Yao [30] stated in 2003 that (because of the integral, fuzzy nature of the RAPD technique that uses multiple random primers,) the selection of the types and quantity of random primers might have a significant influence on the detection of differences in RAPD molecular marker polymorphisms. These authors highlighted the possible threat to the representativeness and objectiveness of RAPD posed by the unbalanced or even biased selection of random primers, when RAPD polymorphism comparisons are used as the macrocosmic tool for molecular systematics. The selection of random primers in RAPD studies, however, has attracted a great deal of scientific attention [16,30,38–45], and RAPD analyses of *C. sinensis* samples have been published by many authors (e.g., Chen et al. [39] published 18 sets of primer data (45%) from of the RAPD results using 40 primers; Feng et al. [40] published 5 sets of primer data (7.4%) from of the RAPD results using 68 primers; Li et al. [41] published 8 sets of primer data (40%) from of the RAPD results using 20 primers; Qian et al. [42] published 19 sets of primer data (63%) from of the RAPD results using 30 primers; Xie et al. [43] published 9 sets of primer data (22.5%) from of the RAPD results using 40 primers; and Yu et al. [44] published 8 sets of primer data (12%) from of the RAPD results using 65 primers). Although there were such selections, these reports truly identified large polymorphic variations among *C. sinensis* samples and other *C. sinensis*-related samples, regardless of how the authors explained their study data or the conclusions drawn from the studies. Consistent with Jiang and Yao [30], Tables 2 and 3 of our study clearly showed large differences in similarities quantified precisely using the ZUNIX equation (2) between sample lanes when using different primers. The most interesting case, however, is the study by Wei et al. [16], who listed only 5 random primers (S23, S33, S61, OPV-14 and OPW-06) and did not discuss the selection criteria used or otherwise demonstrate the representativeness and objectiveness of their

selection. Of the 10 random primers used in our study (*cf.* Table 1) to test the different algorithms (further tests using additional primers are ongoing), the 5 used by Wei et al. [16] showed less variable RAPD polymorphisms between the *H. sinensis* and *C. sinensis* samples. The other 5 primers appeared to demonstrate much greater differences in RAPD polymorphisms between the *H. sinensis* and *C. sinensis* samples, regardless of whether the density-unweighted or weighted algorithm was used (data not shown). In fact, the primers (S7 and S31) shown in Figures 3&4 were not selected by Wei et al. [16], even though they have been used and published by other authors [41–42,44]. Our data, combined with the data in the literature, suggested the following for a fair study design and objective RAPD analysis of *C. sinensis*: (i) the balanced and unbiased selection (the type and quantity) of random primers (this was also suggested by Jiang and Yao [30] in 2003), and (ii) the capture of quantitative molecular information (densities of all DNA amplicons) for similarity computation and cluster construction.

The criteria for the selection of random primers depend on the number of PCR DNA bands and the clarity of the high-density bands [38–45]. However, these criteria may be biased by heavier weighting toward those fungi with relatively stronger expression at a particular stage of *C. sinensis* maturation and would be unfavorable to those fungi with weaker expression at that stage. Considering that *C. sinensis* contains multiple intrinsic fungi from over 30 genera and at least 6 genotypes of *O. sinensis* that are expressed at different levels in different compartments of *C. sinensis* during maturation and that quantitative analytical means are available [6-28,47], mycology researchers may be advised to reconsider their study design (including criteria for the selection of random primers) for RAPD studies of *C. sinensis*, to avoid the selective and unbalanced use of only a few random primers. In fact, RAPD primers that amplify fewer DNA bands may actually reflect a

unique feature or special genetic sequences of an intrinsic fungus or a group of fungi. Thus, screening out RAPD results of this type will probably unlikely result in a fair, balanced conclusion.

To present a global profile for the genetic distance and evolution of *O. sinensis*, scholars have performed integrated analyses of multiple RAPD gels in previous RAPD studies [16,38–45]. To avoid unbalanced weighting of the molecular information represented by the DNA band density in multiple RAPD gels when using new density-weighted algorithms to uncover macrocosmic quantitative molecular information, across-gel normalization must be implemented prior to integral similarity computation and cluster tree construction. Because of our incomplete knowledge of the genomes of the entire intrinsic fungal community in the integrated micro-ecosystem of *C. sinensis*, particularly about the non-culturable fungi that may account for the vast majority of the intrinsic fungi and that contain certain special DNA structures/conformation in the noncoding areas, the quantitative measurement of all intrinsic fungi and their ratios with internal controls has not been possible. In addition, *C. sinensis* maturation has been found to be associated with large, dynamic changes in the differential expression of certain fungi or their mutant genotypes, resulting in an even more complicated background of the intrinsic fungal community [18–19,21,24–28]. Therefore, RAPD functioning as a macrocosmic, fuzzy molecular technique must be implemented quantitatively. In addition to the quantitative handling of the genomic DNA preparation and PCR portions of the experiments, the accurate loading of a molecular weight standard (DNA ladder) onto the agarose gels and the total density quantification of the standard were the key for the across-gel normalization. Such normalization provides a weighted background for the rough measure of DNA template copy number, relatively reflecting the ratios between different intrinsic fungi in *C. sinensis*.

Extended RAPD experiments have been conducted with the use of additional random primers for the unbiased and systematic study of *C. sinensis* [49], where the new algorithms described in this methods paper revealed changes in molecular marker polymorphisms expressed in the caterpillar body, stroma and ascocarp portion of *C. sinensis*. We propose that our findings reflect the dynamic changes in the expression of the multiple intrinsic fungi and fungal genotypes in the different compartments of *C. sinensis* under the multiple fungi-integrated micro-ecosystem hypothesis [17-19,21,23-28,30,33,47], rather than genetic variations under the single-fungus hypothesis [16,29]. Clearly, our interpretation contradicts the hypothesis that the evolutionary distances of *O. sinensis* are responsible for the differences in the *C. sinensis* populations in different production areas, which has been proposed by some *C. sinensis* researchers [38-39,43,45]. We also showed that the RAPD molecular marker polymorphisms of the ascocarp portion are different from those of *H. sinensis*, indicating differences in the fungal genetic background. This is consistent with the findings shown on the RAPD gel images published by Li et al. [41], regardless of the percentage similarity that was computed and published by those authors. A mass spectrum study on proteomic polymorphisms of natural *C. sinensis* has also been performed in our lab using density-weighted algorithms for similarity computation and cluster construction, revealing dynamic changes in proteomic profiles in the different parts of *C. sinensis* at different maturation stages (manuscript revision submitted). These large differences and dynamic changes in RAPD molecular marker and proteomic polymorphisms are also consistent with the microscopic findings of microcycle conidiation of ascospores by Liu et al. [50]: “a similar, but slightly different” morphology of the conidiogenous structures of the microcycle conidiation of ascospores compared to the conidial fructification of *H. sinensis*. It was also

observed that microcycle conidiation of ascospores and microcycle conidiation of conidia produced the conidia with different microscopic shapes (oval vs. sphere) and sizes (5-10 μ m X 4-5 μ m vs. 1 μ m X 1 μ m) [51]. No genotyping information to date has been reported for the different conidiogenous structures or for the different conidia discovered in microcycle conidiation of ascospores/conidia studies. Both the macrocosmic RAPD global profiling and microscopic sequencing and morphological findings challenge the hypothesis that *C. sinensis* is one species of fungus [16,29], but support the hypothesis that *C. sinensis* is an integrated micro-ecosystem with multiple intrinsic fungi from over 30 genera [17-19,21-28,30,33,47]. More and more mycologists have realized that symbiotic interactions between multiple intrinsic fungi in *C. sinensis* may be the key to the induction of the *C. sinensis* sexual fruiting body and formation of ascospores, both naturally and in artificial settings, in strict adherence to Koch’s postulates. The new algorithms and analytical methodologies presented here provide scientists with accurate macrocosmic analytical means to trace the dynamic changes in the molecular marker polymorphism profiles in *C. sinensis*.

In conclusion, the new density-weighted algorithmic methods represent an advance in similarity computation and cluster construction when macrocosmically comparing polymorphisms of RAPD molecular markers in micro-ecosystems for molecular systematics and provide accurate analytic tools to capture the molecular information related to the dynamic changes in differential expression of *C. sinensis* intrinsic fungi during maturation. The apparent dissimilarity in RAPD polymorphisms between *H. sinensis* and *C. sinensis* suggests differences in fungal genetic background and challenges the hypothesis of the anamorph-teleomorph connection between *H. sinensis* and *C. sinensis*.

Acknowledgments

The authors are grateful to Prof. YL Guo of Institute of Microbiology, Chinese Academy of Sciences for providing the *H. sinensis* strain, and to Prof. PY Xu, Ms. M Yang, Mr. W Chen, Mr. JJ Li, and Mr. YC Zhou for their assistance during purchasing and transporting the wild *C. sinensis*.

References

1. Xu, J.T. *Chinese Medicinal Mycology*; Beijing Medical University Press and China Peking Union Medical University Press: Beijing, 1997; pp354-385.
2. Zhu, J.-S.; Halpern, G.M.; Jones, K. The scientific rediscovery of a precious ancient Chinese herbal regimen: *Cordyceps sinensis*: Part I. *J Altern Complem Med* 1998a, 4(3), 289-303. DOI: [10.1089/acm.1998.4.3-289](https://doi.org/10.1089/acm.1998.4.3-289)
3. Zhu, J.-S.; Halpern, G.M.; Jones, K. The scientific rediscovery of an ancient Chinese herbal medicine: *Cordyceps sinensis*: Part II. *J Altern Complem Med*, 1998b, 4(4), 429-457. DOI: [10.1089/acm.1998.4.429](https://doi.org/10.1089/acm.1998.4.429)
4. Tan, N.Z.; Berger, J.L.; Zhang, Y.; Prolla, T.A.; Weindruch, R.; Zhao, C.S.; Bartlett, M.; Zhu J-S. The lifespan-prolonging effect of *Cordyceps sinensis* Cs-4 in normal mice and its molecular mechanisms. *FASEB J*, 2011, 25(1), 599.1
5. Sung, G.H.; Hywel-Jones, N.L.; Sung, J.M. Luangsa-ard, J.J.; Shrestha, B.; Spatafora, J.W. Phylogenetic classification of Cordyceps and the clavicipitaceous fungi. *Stud Mycol*, 2007, 57(1), 5-59. DOI: [10.3114/sim.2007.57.01](https://doi.org/10.3114/sim.2007.57.01)
6. Zhao, J.; Wang, N.; Chen, Y.Q.; Li, T.H.; Qu, L.H. Molecular identification for the asexual stage of *Cordyceps sinensis*. *Acta Sci Natur Univ Sunyatseni*, 1999, 38(1), 122-123.
7. Wang, N.; Chen, Y.Q.; Zhang, W.M.; Li, T.H.; Qu, L.H. Molecular evidences indicating multiple origins in the entomogenous *Cordyceps*. *Acta Sci Natur Univ Sunyatseni*, 2000, 39(4), 70-73.
8. Chen, Y.Q.; Wang, N.; Qu, L.H.; Li, T.H.; Zhang, W.M. Determination of the anamorph of *Cordyceps sinensis* inferred from the analysis of the ribosomal DNA internal transcribed spacers and 5.8S rDNA. *Biochem Syst Ecol*, 2001, 29, 597-607. DOI: [10.1016/S0305-1978\(00\)00100-9](https://doi.org/10.1016/S0305-1978(00)00100-9).
9. Kinjo, N.; Zang, M. Morphological and phylogenetic studies on *Cordyceps sinensis* distributed in southwestern China. *Mycoscience*, 2001, 42, 567-574. DOI: [10.1007/BF02460956](https://doi.org/10.1007/BF02460956)
10. Liu, Z.Y.; Yao, Y.J.; Liang, Z.Q.; Liu, A.Y.; Pegler, D.N.; Liu, A.Y.; Pegler, D.N.; Chase, M.W. Molecular evidence for the anamorph-teleomorph connection in *Cordyceps sinensis*. *Mycol Res*, 2001, 105(7), 827-832. DOI: [10.1017/S095375620100377X](https://doi.org/10.1017/S095375620100377X)
11. Chen, Y.Q.; Wang, N.; Zhou, H.; Qu, L.H. Differentiation of medicinal *Cordyceps* species by rDNA ITS sequence analysis. *Planta Medica*, 2002, 68, 635-639. DOI: [10.1055/s-2002-32892](https://doi.org/10.1055/s-2002-32892)
12. Chen, Y.Q.; Hu B, Xu, F.; Zhang, W.M.; Zhou, H.; Zhou, H.; Qu, L.H. Genetic variation of *Cordyceps sinensis*, a fruit-body-producing entomopathogenic species from different geographical regions in China. *FEMS Microbiol Lett*, 2004, 230, 153-158. DOI: [10.1016/S0378-1097\(03\)00889-9/full](https://doi.org/10.1016/S0378-1097(03)00889-9/full)
13. Chen, J.; Zhang, W.; Lu, T. Morphological and genetic characterization of a cultivated *Cordyceps sinensis* fungus and its polysaccharide component possessing antioxidant property in H22 tumor-bearing mice. *Life Sci*, 2006, 78(23), 2742-2748. DOI: [10.1016/j.lfs.2005.10.047](https://doi.org/10.1016/j.lfs.2005.10.047)
14. Jiang, Y.; Yao, Y.J. ITS sequence analysis and ascoma development of *Pseudogymnoascus roseus*. *Mycotaxon*, 2005, 94, 55-73.
15. Leung, P.H.; Zhang, Q.X.; Wu, J.Y. Mycelium cultivation, chemical composition and antitumour activity of a *Tolyptocladium* sp. fungus isolated from wild *Cordyceps sinensis*. *J Appl Microbiol*, 2006, 101(2), 275-283. DOI: [10.1111/j.1365-2672.2006.02930.x](https://doi.org/10.1111/j.1365-2672.2006.02930.x)
16. Wei, X.L.; Yin, X.C.; Guo, Y.L.; Shen, N.Y.; Wei, J.C. Analyses of molecular systematics on *Cordyceps sinensis* and its related taxa.

- Mycosystema*, 2006, 5(2), 192-202.
17. Stensrud, Ø.; Schumacher, T.; Shalchian-Tabrizi, K.; Svegardenib, I.B.; Kausrud, H. Accelerated nrDNA evolution and profound AT bias in the medicinal fungus *Cordyceps sinensis*. *Mycol Res*, 2007, 111, 409-415. DOI: [10.1016/j.mycres.2007.01.015](https://doi.org/10.1016/j.mycres.2007.01.015)
 18. Zhu, J.-S.; Guo, Y.L.; Yao, Y.S.; Zhou, Y.J.; Lu, J.H.; Qi, Y.; Liu, X.J.; Wu, Z.; Chen, W.; Zhang, L.; Yin, W.T.; Zheng, T.Y.; Zhang, L.J. Maturation of *Cordyceps sinensis* associates with co-existence of *Hirsutella sinensis* and *Paecilomyces hepiali* DNA and dynamic changes in fungal competitive proliferation predominance and chemical profiles. *J. Fungal Res*. 2007, 5(4), 214-224. DOI: [10.3969/j.issn.1672-3538.2007.04.009](https://doi.org/10.3969/j.issn.1672-3538.2007.04.009).
 19. Yang, J.L.; Xiao, W.; He, H.H.; Zhu, H.X.; Wang, S.F.; Cheng, K.D.; Zhu, P. Molecular phylogenetic analysis of *Paecilomyces hepiali* and *Cordyceps sinensis*. *Acta Pharmaceutica Sinica*, 2008, 43(4), 421-426. DOI: [10.1111/j.1745-7254.2008.00754.x](https://doi.org/10.1111/j.1745-7254.2008.00754.x)
 20. Hao, J.J.; Cheng, A.; Liang, H.H.; Yang, X.L.; Li, A.; Zhou, T.S.; Zhang, W.J.; Chen, J.K. Genetic differentiation and distributing pattern of *Cordyceps sinensis* in China revealed by rDNA ITS sequences. *Chin Tradit Herbal Drugs*, 2009, 40(1), 112-116.
 21. Xiao, W.; Yang, J.P.; Zhu, P.; Cheng, K.D.; He, H.X.; Zhu, H.X.; Wang, Q. Non-support of species complex hypothesis of *Cordyceps sinensis* by targeted rDNA-ITS sequence analysis. *Mycosystema*, 2009, 28(6), 724-730.
 22. He, S.Q.; Jin, X.L.; Luo, J.C.; Wang, C.M.; Wang, S.R. Morphological characteristics and submerged culture medium screening of a strain of *Geomyces pannorum* isolation from *Cordyceps sinensis*. *Chin J Grassland*, 2010, 32(z.), 70-75.
 23. Zhang, Y.J.; Sun, B.D.; Zhang, S.; Wangmu; Liu, X.Z.; Gong, W.F. Mycobiota investigation of natural *Ophiocordyceps sinensis* based on culture-dependent investigation. *Mycosystema*, 2010, 9(4), 518-527.
 24. Zhu, J.-S.; Gao, L.; Li, X.H.; Yao, Y.S.; Zhou, Y.J.; Zhao, J.Q.; Lu, J.H. Maturation alterations of oppositely orientated rDNA and differential proliferations of CG:AT-biased genotypes of *Cordyceps sinensis* fungi and *Paecilomyces hepiali* in natural *Cordyceps sinensis*. *Am J Biomed Sci*, 2010, 2(3), 217-238. doi: [10.5099/aj100300217](https://doi.org/10.5099/aj100300217).
 25. Gao, L.; Li, X.H.; Zhao, J.Q.; Lu, J.H.; Zhu, J.-S. Detection of multiple *Ophiocordyceps sinensis* mutants in premature stroma of *Cordyceps sinensis* by MassARRAY SNP MALDI-TOF mass spectrum genotyping. *Beijing Da Xue Xue Bao*, 2011, 43(2), 259-266.
 26. Gao, L.; Li, X.H.; Zhao, J.Q.; Lu, J.H.; Zhu, J.-S. Maturation of *Cordyceps sinensis* associates with alterations of fungal expressions of multiple *Ophiocordyceps sinensis* mutants with transition and transversion point mutations in stroma of *Cordyceps sinensis*. *Beijing Da Xue Xue Bao*, 2012, 44(3), 454-463.
 27. Yao, Y.S.; Zhou, Y.J.; Gao, L.; Lu, J.H.; Zhu, J.-S. Dynamic alterations of the differential fungal expressions of *Ophiocordyceps sinensis* and its mutant genotypes in stroma and caterpillar during maturation of natural *Cordyceps sinensis*. *J Fungal Res*, 2011, 9(1), 37-49,53.
 28. Zhu, J.-S.; Zhao, J.G.; Gao, L.; Li, X.H.; Zhao, J.Q.; Lu, J.H. Dynamically altered expressions of at least 6 *Ophiocordyceps sinensis* mutants in the stroma of *Cordyceps sinensis*. *J Fungal Res*, 2012, 10(2), 100-112.
 29. Guo, Y.L.; Xiao, P.G.; Wei, J.C. On the biology and sustainable utilization of the Chinese Medicine treasure *Ophiocordyceps sinensis* (Berk.) G. H. Sung et al. *Modern Chin Med*, 2010, 12(11), 3-8.
 30. Jiang, Y.; Yao, Y.J. A review for the debating studies on the anamorph of *Cordyceps sinensis*. *Mycosystema*, 2003, 22(1), 161-176.
 31. Jin, Z.X.; Yang, S.H. Progresses and trends of *Cordyceps sinensis* studies. *J Tianjin Med Univ*, 2005, 11(1), 137-140.
 32. Liang, Z.Q. Anamorphs of *Cordyceps* and their determination. *Southwest Chin J Agricultur Sci*, 1991, 4(4), 1-8.
 33. Liang, Z.Q.; Han, Y.F.; Liang, J.D.; Dong, X.; Du, W. Issues of concern in the studies of *Ophiocordyceps sinensis*. *Microbiol Chin*,

- 2010, 37(11), 1692-1697.
34. Liu, F.; Wu, X.L.; Chen, S.J.; Yin, D.H.; Zeng, W. Zhong, G.Y. Advances in studies on artificial culture of *Cordyceps sinensis*. *Chin Tradit Herbal Drugs*, 2007, 38(2), 302-305.
 35. Shen, N.Y.; Zheng, L.; Zhang, X.C.; Wei, S.L.; Zhou, Z.R.; San, Z. Anamorph of *Cordyceps sinensis* (Berk) Sacc. in Monograph for *Cordyceps* Studies 1980-1985. Qinghai Acad Animal Sci Veterin Med: Xining Qinghai, 1983; pp1-13.
 36. Yu, Y.X. Studies on artificial culture of *Cordyceps sinensis*. *J Fungal Res*, 2004, 2(2), 42-46.
 37. Ma, S.L.; Li, Y.L.; Xu, H.F.; Zhang, Z.H. Liu, X. Analyzing bacterial community in young *Hepialus* of intestinal tract of *Cordyceps sinensis* in Qinghai Province. *Chin J Grassland*, 2010, 32(z.), 63-65.
 38. Chen, Y.J.; Wang, W.; Yang, Y.X.; Su, B.; Zhang, Y.P.; Xiong, L.Y.; He, Z.F.; Shu, C.; Yang, D. Genetic Divergence of *Cordyceps sinensis* as Estimate by Random Amplified Polymorphic DNA Analyses. *Acta Genet Sinica*, 1997, 24(5), 410-416.
 39. Chen, Y.J.; Zhang, Y.P.; Yang, Y.X.; Yang, D.R. Genetic diversity and taxonomic implication of *Cordyceps sinensis* as revealed by RAPD markers. *Biochem Genet*, 1999, 37, 201-213. DOI: [10.1023/A:1018738706123](https://doi.org/10.1023/A:1018738706123)
 40. Feng, K.; Wang, S.; Hu, D.J.; Yang, F.Q.; Wang, H.X.; Li, S.P. Random amplified polymorphic DNA (RAPD) analysis and the nucleosides assessment of fungal strains isolated from natural *Cordyceps sinensis*. *J Pharmaceut Biomed Anal*, 2009, 50, 522-526. DOI: [10.1016/j.jpba.2009.04.029](https://doi.org/10.1016/j.jpba.2009.04.029)
 41. Li, Z.Z.; Huang, B.; Li, C.R.; Fan, M.Z. Molecular evidence for anamorph determination of *Cordyceps sinensis* (Berk.) Sacc. *Mycosystema*, 2001, 9(1), 60-64.
 42. Qian, M.; Zeng, W.; Zhang, D.L. RAPD polymorphism analysis on *Cordyceps sinensis*. *LiShiZhen Med Materia Medica Res*, 2011, 22(7), 1738-1739. DOI: [10.3969/j.issn.1008-0805.2011.07.090](https://doi.org/10.3969/j.issn.1008-0805.2011.07.090)
 43. Xie, F.; Zhu, Z.X.; Chen, J.J. RAPD and analysis of ITS sequences of *Ophiocordyceps sinensis* in Gansu. *Chin J Microbiol*, 2011, 23(3), 219-222, 229.
 44. Yu, C.; He, J.L.; Zeng, W.; Deng, B. Fingerprint comparison on *Cordyceps sinensis* and artificial ones by RAPD. *Chin Tradit Herbal Drugs*, 2005, 36(2), 274-277.
 45. Zhang, Y.W.; Chen, Y.J.; Shen, F.R.; Yang, Y.X.; Yang, D.R.; Zhang, Y.P. Study of genetic divergence in *Cordyceps sinensis* and *C. crassispors* from Northwest of Yunnan by using RAPD. *Mycosystema*, 1999, 18(2), 176-183.
 46. Nei, M.; Li, W.H. Mathematical model for studying genetical variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA*, 1979, 74(10), 5269-5273.
 47. Zhang, Y.; Zhang, S.; Wang, M.; Bai, F.; Liu, X. High Diversity of the Fungal Community Structure in Naturally-Occurring *Ophiocordyceps sinensis*. *PLoS ONE*, 2010, 5(12), e15570. DOI: [10.1371/journal.pone.0015570](https://doi.org/10.1371/journal.pone.0015570).
 48. Hu, X.; Zhang, Y.J.; Xiao, G.H.; Zheng, P.; Xia, Y.L.; Zhang X.Y.; St. Leger, R.J.; Liu, X.Z.; Wang, C.S. Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus. *Chin Sci Bull*, 2013, 58, 2846-2854. DOI: [10.1007/s11434-013-5929-5](https://doi.org/10.1007/s11434-013-5929-5)
 49. Yao, Y.S.; Gao, L.; Li, Y.L.; Ma, S.L.; Wu, Z.M.; Tan, N.Z.; Wu, J.Y.; Ni, L.Q.; Zhu, J.-S. Amplicon density-weighted algorithms analyze dissimilarity and dynamic alterations of RAPD polymorphisms in the integrated micro-ecosystem *Cordyceps sinensis*. *Beijing Da Xue Xue Bao* 2014 (in press)
 50. Liu, Z.Y.; Liang, Z.Q.; Liu, A.Y. Investigation on microcycle conidiation of ascospores and conidiogenous structures of anamorph of *Cordyceps sinensis*. *Guizhou Agricultural. Sci*, 2003, 31(1), 3-5.
 51. Xiao, Y.Y.; Chen, C.; Dong, J.F. Morphological observation of ascospores of *Ophiocordyceps sinensis* and its anamorph in growth process. *J Anhui Agricult Univ*, 2011, 38(4), 587-591. DOI: [10.1162/S.20110624.1628.030](https://doi.org/10.1162/S.20110624.1628.030)