PeerJ Preprints NOT PEER-REVIEWED

Genetic characterization of *Cycas panzhihuaensis* (Cycadaceae): crisis lurks behind a seemingly bright prospect

Siyue Xiao [1,2], **Yunheng Ji** [1], **Jian Liu** [1], **Xun Gong** Corresp. [1]

[1] Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan province, China
[2] University of Chinese Academy of Sciences, Beijing, China

Corresponding Author: Xun Gong
Email address: gongxun@mail.kib.ac.cn

Background

Cycas panzhihuaensis L. Zhou & S. Y. Yang (Cycadaceae) is an endangered gymnosperm species endemic in the dry-hot valley of Jinsha River basin from southwest China. Although the wild *C. panzhihuaensis* population from Panzhihua Cycad Natural Reserve is well protected, other known populations that fall outside the natural reserve may preserve specific genetic resources while face with larger extinction risk because of lacking essential monitoring.

Methods

In this study, we analyzed the genetic diversity, phylogeographical structure and demographic history of *C. panzhihuaensis* from seven known locations so far by sequencing three chloroplastic DNA regions (*psb*A-*trn*H, *psb*M-*trn*D, and *trn*S-*trn*G), four single-copy nuclear genes (*PHYP*, *AC5*, *HSP70*, and *AAT*) from 61 individuals, and eleven microsatellite loci (SSR) from 102 individuals.

Results and Discussion

We found relative high genetic diversity within populations and high genetic differentiation among the populations of *C. panzhihuaensis*, which is similar with the other Asian inland cycads. Despite no significant phylogeographical structure was detected, small and unprotected populations possess higher genetic diversity and more unique haplotypes, which deserve due attention. Results of demographic dynamics suggest that human activity is the key factor that leads *C. panzhihuaensis* to endangered status. Basing on the genetic characterization of *C. panzhihuaensis*, we proposed several practical guidelines for the conservation of this species, especially for its small populations.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27265v1 | CC BY 4.0 Open Access | rec: 8 Oct 2018, publ: 8 Oct 2018

PeerJ Preprints NOT PEER-REVIEWED

Genetic characterization of *Cycas panzhihuaensis* (Cycadaceae): crisis lurks behind a seemingly bright prospect

Siyue Xiao[1, 2], Yunheng Ji[1], Jian Liu[1], Xun Gong[1]

[1] Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan province, China

[2] University of Chinese Academy of Sciences, Beijing, China

Corresponding Author:

Xun Gong[1]

No. 132 Lanhei RD, Panlong District, Kunming, Yunnan province, 650201, China

Email address: gongxun@mail.kib.ac.cn

ABSTRACT

Background

Cycas panzhihuaensis L. Zhou & S. Y. Yang (Cycadaceae) is an endangered gymnosperm species endemic in the dry-hot valley of Jinsha River basin from southwest China. Although the wild *C. panzhihuaensis* population from Panzhihua Cycad Natural Reserve is well protected, other known populations that fall outside the natural reserve may preserve specific genetic resources while face with larger extinction risk because of lacking essential monitoring.

Methods

In this study, we analyzed the genetic diversity, phylogeographical structure and demographic history of *C. panzhihuaensis* from seven known locations so far by sequencing three chloroplastic DNA regions (*psb*A-*trn*H, *psb*M-*trn*D, and *trn*S-*trn*G), four single-copy nuclear

genes (*PHYP*, *AC5*, *HSP70*, and *AAT*) from 61 individuals, and eleven microsatellite loci (SSR) from 102 individuals.

Results and Discussion

We found relative high genetic diversity within populations and high genetic differentiation among the populations of *C. panzhihuaensis*, which is similar with the other Asian inland cycads. Despite no significant phylogeographical structure was detected, small and unprotected populations possess higher genetic diversity and more unique haplotypes, which deserve due attention. Results of demographic dynamics suggest that human activity is the key factor that leads *C. panzhihuaensis* to endangered status. Basing on the genetic characterization of *C. panzhihuaensis*, we proposed several practical guidelines for the conservation of this species, especially for its small populations.

INTRODUCTION

The protection of genetic diversity within species has become one of the main objectives of conservation efforts (Hamrick & Godt 1996; Rauch & Bar-Yam 2004; Pierson et al. 2016). Meanwhile, the analysis of population genetics has been widely applied in species conservation and management (Pérez-Espona 2017; Nash et al. 2018). Discrete management of populations with distinct genetic backgrounds can preserve the natural genetic patterns of species, avoid inbreeding and outbreeding depression and increase species-wide resilience (Schindler et al. 2010; Larson et al. 2014). Genetic analysis based on molecular data in recent decades provide the guidance to us for plant resources collection of endangered plants and improve the success of population reintroduction and reinforcement (Griffith et al. 2015; Kaulfuss & Reisch 2017). Combining with historical geographical and climate changes, the genetic data can also be used to infer the population demographical dynamic which helps us in understanding the current extent of a species facing habitat loss, landscapes and climate changes (Selwood et al. 2015).

Cycas panzhihuaensis L. Zhou & S. Y. Yang is a gymnosperm species from the family Cycadaceae. Previous studies considered that *Cycas* represented most ancient lineage ascending to 250 million years ago (Ma) (Hill et al. 2003b), while recent dating results found they were much younger (~12 Ma) (Nagalingum et al. 2011). For cycad, however, species could not be

rescued from endangerment by rapid diversification and phylogenetic diversity. They are facing high risk of losing significant component of evolutionary diversity in the current extinction crisis (Yessoufou et al. 2017). Integrative approach combining multidimensional diversity analyses with biogeographical framework is being increasingly advocated as insufficiency for conservation from whether ancient refuges or protection area currently (Fjeldsa & Lovett 1997; Jetz et al. 2004; Mouillot et al. 2016).

Cycas panzhihuaensis is the only species from *Cycas* sect. *Panzhihuaenses* (D.Yue Wang) K.D.Hill, which is sister with other two coastal species from *C*. sect. *Asiorientales* Schuster and seems to diverge earliest from *Cycas* (Liu et al. 2018). Unlike to the two coastal species, *C. panzhihuaensis*is is endemic in the xerothermic valley along Jinsha River, southwest China, a basal zone adjacent to the Hengduan Mountains Region. The vulnerable valley habitat with dry-hot ecosystems is inhabited by high richness of endemic and endangered plant species (Zhao & Gong 2015), which may be related to its role as refugium in the quaternary glacial period. Investigations in the phylogeographical and demographic history of *C. panzhihuaensis* facilitate the elucidation of the role of geological, climatic and evolutionary factors in shaping the population history of *Cycas* species.

As the most threatened group in gymnosperms, 88% species of cycads (IUCN 2017) are threatened with extinction, which mainly due to human activities including habitat disturbance, over-collecting for ornamental and medical purposes (Mankga & Yessoufou 2017). *Cycas panzhihuaensis* is vulnerable and has suffered serious declining in previous 40 years for illegal trade since it was described in 1981 (Zhou et al. 1981), as the cycads became popular in horticulture since 1970s in China. Combined with frequently deforestation, grazing, agricultural reclamation and mining, half of primitive habitat was decreased (45.22% estimated by He & Li 1999, 20%-50% by Hill et al. 2003a). He & Li (1999) found that only five of the 13 populations remained by 1999. Through a comprehensive investigation along the Jinsha River basin, *C. panzhihuaensis* was mainly found to restrict in the Sichuan Panzhihua Cycad Natural Reserve, which the census population size (*N*c) is large to 240,000, with other isolated populations only accounting for about 1% individuals (Hao 2011).

Several studies have conducted on *C. panzhihuaensis* previously including its geographical distribution (He & Li 1999), habits, morphological characteristics (Liu et al. 1990), breeding (Liu et al. 2016) and community characteristics (Hao 2011; Wang 2016) as well as the whole chloroplast (cp) genome sequencing and characterization (Han et al. 2016). The genetic diversity of this species has also been reported in a series of studies. Relative high genetic variation (percentage of polymorphic loci %P = 90.38%) was found in population Panzhihua Natural Reserve using 13 inter-simple sequence repeat (ISSR) markers (Wang et al. 2010). Subpopulation structure in this reserve was explored and two evolutionary significant units was defined via amplified fragment length polymorphism (AFLP) (Yang et al. 2015). Preliminary genetic characteristics of seven populations were investigated in a previous study by two chloroplastic DNA markers (Zhang, 2012). However, previous population genetic studies are either based on limited populations (Li et al. 1999; Wang et al. 2010; Yang et al. 2015) or insufficient genetic data (Zhang, 2012), and an overall and more precise population genetic study of *C. panzhihuaensis* based on different datasets is still urgent.

In this study, we conducted a comprehensive population genetic study on *C. panzhihuaensis* by using three chloroplast DNA (cpDNA) intergenic spacers, three nuclear DNA (nDNA) genes, and 11 nuclear microsatellite markers to investigate the genetic diversity, genetic structure and population dynamics in *C. panzhihuaensis*. On the basis of such analyses, we are aiming to answer the following questions:

1. What is the general genetic diversity level of *C. panzhihuaensis*?
2. What is the possible dispersal pattern of *C. panzhihuaensis* and what are the factors affecting its dispersal, thus laying a foundation for further planning and reintroducing action?
3. Which populations of *C. panzhihuaensis* need to be particularly focused and protected?

120

121

122

123 ## MATERIALS & METHODS

124 **Sampling**

125 All individuals in this study were collected from seven locations in Yunnan and Sichuan
126 province of China, including five wild populations and two in cultivation, (JPD and GH), as the
127 nearby habitats were completely destroyed. In total, we sampled 102 adult individuals for
128 microsatellite analysis and 5-10 individuals per population for chloroplast and nuclear DNA
129 sequencing. Details of populations and sample size (n) are shown in Table 1.

130

131 **DNA extraction and PCR amplification**

132 Total DNA was extracted from the silica-dried and -20°C stored leaves by modified CTAB
133 method (Doyle 1991). We screened three cpDNA (*psb*A-*trn*H, *psb*M-*trn*D, and *trn*S-*trn*G)
134 (Shaw et al. 2005; Zhang 2012), four nDNA (*PHYP*: phytochrome P gene, *AC5*: *Cycas revoluta*
135 actin 5, *HSP70*: heat shock protein 70 kDa gene and *AAT*: aspartate aminotransferase) (Gong &
136 Gong 2016; Liu 2016), and 11 nuclear SSR markers (Wang et al. 2008; Yang et al. 2008; Li et
137 al. 2009; Zhang et al. 2010) with high polymorphism for subsequent analyses. Table S1 and
138 Table S2 list the description of primers.

139

140 For nDNA, the PCR reaction mix contained 3.0 μL 10×PCR buffer, 2.25 μL $MgCl_2$, 2.25 μL
141 dNTP, 2.25 μL DMSO/BSA, 0.6 μL of each primer, 0.6 μL Taq DNA polymerase (5 U/μL;
142 TsingKe Biological Technology, Kunming, China), 10-50 ng DNA and replenished double-
143 distilled water to a total volume of 30 μL. PCR sequence amplifications for nDNA were an initial
144 5 minutes (min) denaturation at 94 °C; 35 cycles of 5 min denaturation at 95 °C, 1 min annealing
145 at 50-55 °C, and 2 min extension at 72 °C; with a final 10 min extension at 72 °C. For cpDNA,
146 the PCR reaction mix contained 3.0 μL 10×PCR buffer, 1.5 μL $MgCl_2$, 1.5 μL dNTP, 1.5 μL
147 DMSO, 0.45 μL of each primer, 0.45 μL Taq DNA polymerase (5 U/μL), 10 ng DNA and add
148 19.5 μL double-distilled water to a total volume of 30 μL. PCR sequence amplifications for
149 cpDNA were initiated with 5 min denaturation at 80 °C; 30 cycles of 5 min denaturation at 95

150 °C, 1 min annealing at 50-55 °C, and 1.5 min extension at 65 °C; with a final 1.5 min extension
151 at 65 °C. The PCR products were purified and sequenced in both directions by TsingKe
152 Biological Technology, Kunming, China. For microsatellite markers, the primers were labeled
153 fluorescently with FAM and then with capillary electrophoresis for genotyping at Shanghai
154 Majorbio Bio-pharm Technology Company Ltd.
155
156 **Data analysis**
157 **Nucleotide sequence**
158 Nucleotide sequences were visually adjusted by SeqMan (Swindell & Plasterer 1997) in
159 DNASTAR biological information package, and aligned with BioEdit v7.0.4.1 (Hall, 1999).
160 After congruence test, the three cpDNA fragments (P=1.00) and three nuclear gene sequences
161 *AC5*, *HSP70*, and *AAT* (P=0.86) were separately combined by PAUP v4.0 (Swofford 2002). The
162 combined nDNA matrix was named as N3 for subsequent analysis and *PHYP* was analyzed
163 separately as this nuclear gene showed incongruent with other genes. To estimate the genetic
164 diversity, the nDNA sequences were phased and the nucleotide diversity (π), haplotype diversity
165 (Hd) (Nei 1987) were calculated using DnaSP v5.0 (Rozas 2009). Total genetic diversity (H_T),
166 genetic diversity within populations (H_S), and the coefficient of genetic differentiation for non-
167 ordered alleles (G_{ST}, N_{ST}) were calculated using Permut v1.0 (Pons & Peti 1996). The genetic
168 variations within and among populations were evaluated by AMOVA in Arlequin v3.11
169 (Schneider et al. 2000). LDNeInterface v1.3.1 (Waples & Do 2008) was used to estimate the
170 effective population size of each population.
171
172 Maximum Parsimony (MP) 50% majority trees and Bayesian trees were constructed to infer the
173 topology of haplotypes with *C. guizhouensis* Lan & R.F. Zou as outgroup. For the MP trees,
174 multi-trees heuristic search was tested with 1000 replicates bootstrap analysis by PAUP v4.0
175 (Swofford 2002). For Bayesian inference implemented in MrBayes v3.2.1 (Fredrik et al. 2012),
176 the Markov chain Monte Carlo (MCMC) was ran for 10^6 generations and trees were sampled
177 every 100 generations, with the first 2500 trees being discarded. The haplotype median-joining
178 network achieving the common ancestor simulation in 95% confidence interval was generated by
179 the software NETWORK v4.2.0.1 (Forster et al. 2007), with indels were treated as single
180 mutation events.

181

182 The Bayesian method implemented in BEAST v1.84 was also used to estimate the divergence
183 time across haplotypes (Drummond et al. 2012), with a strict clock model option and a Birth-
184 Death Process tree model prior. The best-fit BIC model in BEAST (HKY, Hasegawa-Kishino-
185 Yano model all applied to cpDNA, *PHYP* and N3) was defined by W-IQ-TREE Modelfinder
186 (Trifinopoulos et al. 2016). The average nucleotide synonymous substitution rate was estimated
187 as 1.01×10^{-9} per site per year mutation for cpDNA, and 6.1×10^{-9} per site per year mutation for
188 nDNA, which is concluded from the average rates of seed plants (Graur & Li 2000). MCMC
189 runs for 10^8 generations and were sampled every 10^4 generations with 25% trees were discarded
190 as burn-in.

191

192 To detect the past population fluctuation events, DnaSP v5.0 was conducted to calculate
193 Tajima's D and Fu and Li's neutrality tests and to generate mismatch distributions plots. The
194 sum of squared deviation (SSD) and Harpending's raggedness (r) of mismatch distributions were
195 calculated in Arlequin v3.11. (Schneider et al. 2000) A Bayesian skyline plot analyses were
196 performed by BEAST v1.84 (Drummond et al. 2012) to infer the historical demography. The
197 effective sample sizes (ESS>200) and convergence to the stationary distribution were examined
198 in Tracer, v1.6 (Rambaut et al. 2014).

199

200 **Microsatellite data analysis**

201 The parameters of genetic diversity and differentiation were calculated by GenAlEx v6.3
202 (Peakall & Smouse 2006), along with Mantel tests which evaluated the isolation by distance
203 (IBD) model using a correlation of geographic distance and genetic distance among every pair of
204 populations. The F_{ST} acquired from AMOVA analysis in Arlequin was used to calculate the gene
205 flow (Nm) via formula $N\text{m}=(1-F_{ST})/4F_{ST}$ (Wright 1931). A Bayesian clustering method was
206 conducted in the STRUCTURE v2.2 (Pritchard et al. 2000) for grouping the populations. Twenty
207 separate runs with group number K from one to seven were performed. Each run was estimated
208 as 10^5 MCMC steps and a 10^5 iterations burn-in. The best fit number of genetic clusters was
209 deduced by ΔK using STRUCTURE HARVESTER (Earl & vonHoldt 2012). Principal
210 coordinate analysis (PCoA) based on individuals was generated by MVSP v3.12 (Kovach 1998).
211 The BOTTLENECK v1.2.02 (Piry et al. 1999) software were performed to explore recent

212 bottlenecks events and Garza-Williamson index (GWI; M-ratio) were used to detect early shrink
213 in population size (Schneider et al. 2000). GWI is the ratio of alleles number to range in allele
214 size, and the lower than 0.68 index suggested a relative early reduction in population size when
215 more than seven loci are analyzed. The GWI was calculated by Arlequin v3.1. We used the
216 linkage disequilibrium (LD) method in LDNeInterface v1.3.1 (Waples & Do 2008) to estimate
217 effective population sizes (*N*e) atthree levels of lowest allele frequency (0.01, 0.02, and 0.05) for
218 a 95% confidence interval.

219 RESULTS

220 **DNA sequences characterization**

221 The combined cpDNA alignment of *psb*A-*trn*H, *psb*M-*trn*D and *trn*S-*trn*G varied from 2634 to
222 2741bp with 137 polymorphic sites were totally detected. Among them, four substitutions and
223 three inserts/indels from combined cpDNA generated 10 haplotypes in seven populations of *C.*
224 *panzhihuaensis*. The cpDNA haplotype Hap1 (cp_H1) is the most widely shared haplotype,
225 which was detected in PDH, PZH, DSS, GH and SL. Haplotype cp_H2 was shared by PZH and
226 SL. The remaining haplotypes were specific to single populations (Table S3, Fig. 1a).
227
228 *PHYP* sequence was aligned with a length of 970-1007bp, containing 42 polymorphic sites, and
229 five haplotypes were derived from the seven populations. The most common haplotype P_H1
230 occurred in all populations. P_H3 was shared by three unprotected populations WQ, SL and GH.
231 P_H2 was private to populations PZH, and P_H4 and P_H5 were only found in SL (Figure 1b).
232 The nuclear gene, *AC5*、*HSP70* and *AAT* were aligned as 862bp, 715-716bp, and 571bp
233 respectively. When combined, the sequences varied from 2148 to 2149bp in length, containing
234 eight polymorphic sites in total. For the combined three nuclear genes, we found 11 haplotypes
235 from seven populations. Among them, N3_H1 and N3_H2 were the most widely distributed
236 haplotypes, which presented in six and four populations respectively. In contrary, N3_H5 and
237 N3_H6 were only found in PZH, N3_H8 merely occurred in WQ, while N3_H9, N3_H10 and
238 N3_H11 were only detected in GH (Figure 1c).
239

240 **Genetic diversity and differentiation**

241 Relative lower total nucleotide diversity (π) and higher haplotype diversity (*Hd*) of *C.*
242 *panzhihuaensis* were detected in combined cpDNA (π = 0.00250, *Hd* = 0.8033) than nDNA
243 (*PHYP*: π = 0.00040, *Hd* = 0.3302; N3: π = 0.00056, *Hd* = 0.7350) (Table 1). Populations WQ,
244 SL and GH displayed haplotype diversity in cpDNA level (Table S3). For nuclear DNA,
245 populations PZH, WQ, SL and GH showed haplotype diversity detected by *PHYP*, and six
246 populations except JPD showed haplotype diversity which was revealed by N3. Results of
247 cpDNA, nDNA and microsatellite were accordant, with three small and unprotected populations
248 WQ, SL and GH displaying higher within-population haplotype diversity, and population SL
249 possesses the most private haplotypes in all datasets.
250
251 The total genetic diversity of *C. panzhihuaensis* inferred by chloroplast and nuclear DNA were
252 higher than the average intrapopulation diversity (cpDNA: H_T =0.881, H_S =0.235; *PHYP*: H_T
253 =0.366, H_S =0.257; N3: H_T =0.769, H_S =0.506) (Table 1), which were consistent with the result
254 of AMOVA analysis (Table 3). For cpDNA, nearly 87% variations existed among populations,
255 with only 13% occurred within populations. For nuclear DNA, the variations among-population
256 and within-population were 32.73% vs 67.27% (*PHYP*) and 40.83% vs 67.27% (N3),
257 respectively. The results of gene flow were 0.039 to 0.514 for all datasets (Table 4).
258
259 **Spatial genetic structure**
260 For the U test, no significant difference was detected between each pair of G_{ST} and N_{ST} (Table
261 1). Meanwhile, no significant correlation between genetic distance and geographical distance
262 (P>0.05) was detected by mantel tests based on all DNA datasets (Fig S1), which indicated a
263 non-significant effect of isolation by distance (IBD) model.
264
265 For the clustering analyses from SSR phenotypes (Fig S3), all individuals obscurely distributed
266 in two-dimensional PCoA plot with no clear groups can be clarified. Meanwhile, STRUCTURE
267 analysis separated the 102 individuals from seven populations into five genetic components,
268 corresponding to DSS, PDH, PZH, WQ and a part of SL in general. Individuals of JPD and
269 partial SL were grouped into PZH, and GH clustered to DSS.
270
271 **Haplotype phylogeny and network analysis**

The network inferred from cpDNA showed a relative closer relationship between private haplotypes (cp_H3-5 from WQ and cp_H6-8 from SL), while they shared distant relationship with more widely shared Hap_1 (Fig 2 a). For *PHYP*, a missing haplotype occurred between the widely shared P_H1, P_H3 and private P_H4-5 which belong to SL (Fig 2 b). For the combined N3, the most widely shared haplotype N3_H1 and N3_H2 located at the center of the network, connecting the other private haplotypes to a ring-and-star manner by one step (Fig 2 c).

Due to insufficient informative sites, haplotypes from cpDNA, *PHYP* and combined N3 diverged parallelly and formed separate comb-like cladograms. The divergent time between *C. panzhihuaensis* and outgroup *C. guizhouensis* estimated by Bayesian implement method were 1.70 Ma, 1.19Ma and 0.18 Ma from cpDNA, *PHYP* and N3 data respectively. Inferences of the divergent time in the crown node of haplotypes were 0.39 Ma, 0.19 Ma and 0.10 Ma from corresponding datasets, and the estimate time among different haplotypes on the internal divergence node were 0.02 – 0.09 Ma (Fig 2). This indicates that the divergence among haplotypes occurred in the mid-late Pleistocene.

Demographic analysis

The Tajima's D and Fu and Li's test based on cpDNA and *PHYP* displayed all positive and negative values respectively, while for N3, negative Tajima's D and Fu's Fs and positive Fu and Li's *F** and Fu and Li's *D** was detected (Table 4). However, all above indices showed non-significance, implying that *C. panzhihuaensis* may have not experienced bottleneck or expansion effect. Mismatch distributions under stable population model displayed multimodal for cpDNA, which was consistent with the neutral test results and indicated the demographic equilibrium of populations (Fig S3). However, flat curve or unimodal was detected from *PHYP* and N3, suggesting the existence of recent expansion. No significance of SSD and raggedness index showed unimodal model could not be rejected.

The Bayesian Skyline Plot of different datasets showed a coherent relative stable (cpDNA and N3) or slight increasing trend (*PHYP*) of population size in history, and a slight recent decline at about 1000 years ago (cpDNA and N3) to hundreds of years ago (*PHYP*) (Fig 3). The results of shifted mode and significantly different T.P.M and S.M.M indicated that population DSS and

303 JPD had experienced severe bottlenecks events (Table 5), which was consistent with the BSP
304 result. Garza-Williamson index (Table 5) under 0.68 means all populations experienced a
305 decrease.

306 CONCLUSIONS and DISCUSSION

307 Genetic diversity and structure

308 When facing with the estimation of nucleotide diversity, the selection of DNA sequences or
309 molecular markers, population sizes, the sampling modes and sample sizes will all make
310 discrepancies (Reisch & Bernhardt-Römermann 2014). Comparing with the former study (π =
311 0.0038; *Hd* = 0.571; F_{ST} = 0.7903, Zhang, 2012) only based on two cpDNA *psb*A-*trn*H and *trn*S-
312 *trn*G but larger sampling size per population, estimates on the genetic diversity and fixation
313 index of *C. panzhihuaensis* in this study (π = 0.00259; *Hd* = 0.8033; F_{ST} = 0.866) are closer to
314 other inland Asian *Cycas* species (π = 0.0009-0.0026; *Hd* =0.492-0.940, aver. *Hd* =0.691; F_{ST} =
315 0.6982-9980, aver. F_{ST} = 0.8690, Zheng, 2017).
316
317 For the SSR analysis, our study sampled more populations, larger sample sizes in contrast
318 to previous studies based on AFLP (Yang et al. 2015) and ISSRs (Wang et al. 2010). As a result,
319 the total percentage of polymorphic loci and expected heterozygosity detected through SSR was
320 lower than AFLP and ISSRs (*%P* = 72.73%, *H*e = 0.328, this study, *%P* = 94.12%, *H*e = 0.335,
321 Yang, 2015; and *%P* = 90.38%, Wang, 2010), also closer to the other inland Asian *Cycas* species
322 inferred from SSR datasets (*%P* = 66.67-94.12, aver. *%P* = 66.67-94.12, *He* = 0.2470-0.5430,
323 Zheng, 2017).
324
325 Yang (2015) concluded *C. panzhihuaensis* possessed more ancestral polymorphism for its higher
326 genetic diversity than other inland species, but we argued there was no significant difference
327 between *C. panzhihuaensis* and other *Cycas* species though *C. panzhihuaensis* relatively
328 diverged earlier than other *Cycas* in evolutionary history and occupied larger *N*c and narrower
329 distribution area. The inland Asian *Cycas*, including *C. panzhihuaensis*, with similar habitat traits
330 as well as morphological, life history and reproductive characteristics were detected as similar
331 genetic traits - relative high level of genetic variation and significant genetic differentiation
332 among populations. Two coastal species *C. revoluta* and *C. taitungensis*, in contrast, with less
333 geographical barrier among populations, possess much greater genetic diversity and less

interpopulation genetic difference (π = 0.0127-0.0581; *Hd* =0.959-0.998; F_{ST} = 0.056-0.864 (Huang et al. 2001; Chiang et al. 2009), even though they were revealed as sister of *C. panzhihuaensis* (Liu et al. 2018). These results are consistent with the hypothesis that genetic diversity is weakly correlated with population and sample size (Reisch & Bernhardt-Römermann 2014), and strongly related to eco-landscape (Reisch et al. 2017). The early isolation of *C. panzhihuaensis* with the two coastal species may have promoted the speciation, but distance did not strongly shape the genetic structure of *C. panzhihuaensis*, which may have been due to various interference of regional environments and invasive species in evolutionary history.

The chloroplast genome of *Cycas* is maternally inherited to the next generation through seed, while nuclear genome is biparental inherited and can reflect the seed flow as well as pollen flows. The fixation index (F_{ST} = 0.331) detected by SSR in our study is higher than the other Asian *Cycas* species (F_{ST} = 0.114-0.296, Zheng, 2017). Estimates of gene flow inferred from DNA datasets were below than critical value 1, indicated a quite low intraspecific genetic flow of *C. panzhihuaensis*. The higher F_{ST} revealed by chloroplast DNA than nuclear DNA and SSR in this study indicated the pollen flow of *C. panzhihuaensis* was much greater than the seed flow. Meanwhile, the dispersion of the sinkable and as large as 2.5cm *Cycas* seed relies on tumbling out naturally. Thus, up to 98.5% seeds drop less than 10 meters of mother plants (Hall & Walter 2011), although in a rare case of *Tupaia belangeri*, also called the northern tree shrew, would account for the dispersal of *C. panzhihuaensis*'s seeds (Yu et al. 2007). Meanwhile, the pollination distance is not far for most of the Cycads (Hall & Walter 2011; Yu 2015), with low transmission efficiency outside 2.55 meters (Wang et al. 1997). We surmise that the low natural fruiting and germination rate of *C. panzhihuaensis* may related to its pollination and seed dispersal barriers.

Demographical history

Unlike those reported inland *Cycas* species (Liu et al. 2015; Zheng et al. 2016; Yang et al. 2017) and the sympatric species distributed in the dry and hot valley of Jinsha River: *Munronia delavayi* (Jia et al. 2014), and *Nouelia insignis* (Gong et al. 2011), no significant spatial genetic structure of *C. panzhihuaensis* was revealed by previous research (Wang et al. 2010; Zhang 2012; Yang et al. 2015) and our study. Most private haplotypes are found in WQ,

SL, GH, which are small populations mostly located in the boundary of the distribution range. This may also attribute to the limited seed flow of *C. panzhihuaensis,* resulting multiple evolution lineages in different isolated locations. As significant demographic contraction was not detected in the Bottleneck analysis and mismatch distribution estimates, and the weak migration ability of *C. panzhihuaensis* reflected from gene flow and AMOVA analysis, the most convincing scenario is that these three small populations has experienced rapid adaptive evolution due to geographical and climatic complexity. Under this hypothesis, the small populations shall be regarded as a source of new endemic species (Kark et al. 1999; Araújo & Williams 2001).

According to Yang, (2015), *C. panzhihuaensis* experienced population growth at c. 0.08 Ma - 1.6 Ma, which covers the haplotype differentiation time obtained in our study (c. 0.1Ma – 0.39 Ma) inferred by Bayesian method. Historical landscape structure may impose more vital influences on genetic diversity than current habitat conditions (Reisch et al. 2017). Southwest China is a known hotspot of biodiversity where the terrain complexity produced unique and diverse climate types (Myers et al. 2000). Jinsha River dry hot valley provided refugia for various surviving species in the Quaternary glaciations (Gong et al. 2011; Zhang et al. 2011a; Zhang et al. 2011b), including the ancestors of *C. panzhihuaensis*.

A plausible scenario is that *C. panzhihuaensis* once widespread and formed an interconnected population at Jinsha River basin in mid-late Pleistocene, given similar dry hot valley did not retain at higher latitudes. In general, the climate oscillation in Pleistocene have greatly influenced *C. panzhihuaensis*. The species experienced multiple population expansion and contraction during the Pleistocene glacial-interglacial cycles (Zachos et al. 2008), causing the inconsistent evolution history of different DNA sequences, which can be verified by the conflicting results of neutrality tests and mismatch analysis in this study. Along with the formation of biogeographic barrier such as landslide dams (Duan et al. 2013), as well as other negative factors including species invasion and limited precipitation, habitat of *C. panzhihuaensis* had been losing and was isolated to several fragmented populations, each of which has experienced the loss of ancestral polymorphism and therefore present the extant

random and unordered genetic structure pattern. Results with each Garza-Williamson index below 0.68 lend support to the populations decline hypothesis.

Furthermore, recent human activities have also threatened the population dynamics of *C. panzhihuaensis*. The results of shifted mode and significantly different T.P.M and S.M.M indicated that population DSS and JPD had experienced severe bottlenecks events (Table 5), which may be the result of recent and rapid colonization on account of artificial or natural selection factors. Results of Bayesian Skyline Plot implied population decline events happened in the past 100 years ago. These results may have direct relationship with long-term habitat loss due to mining and grazing activities and over-exploitation in recent 40 years, which causing severe habitat fragmentation, followed by great genetic diversity loses of *C. panzhihuaensis*.

Implications for conservation

Although the phylogeographical pattern of *C. panzhihuaensis* could not be clearly revealed because the great loss of population size in populations WQ, JPD, SL and GH, we evaluated the overall genetic diversity of *C. panzhihuaensis* within and among the current populations in different levels (cpDNA, nDNA and SSR), which can guide us in the conservation practices. Genetic diversity is not often found to be correlated with the population size (James et al. 2018), which is also consistent with our study. Population PZH seems to be the center of refugium for its considerable numbers of *N*c. However, population WQ and GH have been detected with haplotypes diversity higher than populations PZH (Table S3), although the distribution areas of the former two population are far less than PZH whether in history or at present. However, small populations of *C. panzhihuaensis* have been suffering from severe exploitation and habitat loss (He & Li 1999). Population SL, for example, while the population with 20 hm^2 area containing 2000 adult individuals in 1979 before large scale of collection, the remaining individuals is lower than 20 (Wang 2016). PDH and PZH are the only two populations containing adult individuals more than 100, but also threatened by deforestation from mining and grazing. The existing six isolated populations except PZH present a fragmented habitat with most of their restricted individuals at a young age, which results the populations in a vulnerable status and will be extirpated at any time by demographic stochasticity, environmental or reproductive disturbance (Lande 1999; Segarra-Moragues et al. 2005).

To successfully achieve the conservation of endangered species, it is often insufficient to raise the abundance of offspring, but to maintain the maximum genetic diversity as evolutionary potential is partially driven by genetic diversity (Etterson 2004; Jump & Penuelas 2005; Li et al. 2012). Estimate value of the effective population size (*N*e) for population PZH is only 17.6 (Table 2), despite its considerable *N*c. The value is lower than but close to the previous study (Mode: 29.369, Median: 30.750, Yang, 2015). Total *N*e of species lower than 50 (Total: 39.0) implied the extant achievement of conservation are far from effective according to the criterion that inbreeding depression could not be prevented over five generations for wild plants when *N*e < 50 (Frankham et al. 2014). As much genetic diversity as possible should be contained in the scope of protection in order to avoid inbreeding completely (as *N*e > 100).

The criteria for selecting priority protected populations must include the uniqueness and diversity of the population, particularly in its allelic composition (Petit et al. 2008). For *C. panzhihuaensis*, given the distinct private haplotypes of SL, PZH, WQ and GH and the estimated genetic clustering obtained by STRUCTURE analyses (Fig S2), five evolutionary significant units (ESUs) combining JPD and PZH, DSS and GH separately were defined and should be protected separately as independent management units (MUs) (Griffith et al. 2015). In situ conservation should be implemented as it can effectively resist against the high risk of genetic erosion to small populations. It is imperative and urgent to collect seeds from these populations to nearby reserves or botanical gardens while ensuring the sustainability in situ conservation (Griffith et al. 2017). Interbreeding between JPD and PZH, DSS and GH is considerable to improve their genetic variation of MUs respectively. To avoid the adverse effects of small populations and thus restore the viability and evolutionary potential, considering the extremely poor population size of populations SL, WQ, GH, JPD and DSS, the suitable habitat should to be evaluated and practical population reconstruction and recovery should be carried out in appropriate localities in further studies.

ACKNOWLEDGEMENTS

The authors thank Fangming Zhang for her assistance with plant sampling; Xiuyan Feng, Rui Yang and Yujuan Zhao for their help with methods guiding.

456

457

458

459

460 ## REFERENCES

461 Araújo MB, Williams PH (2001) The bias of complementarity hotspots toward marginal
462 populations. Conserv Biol 15:1710-1720
463 Chiang YC, Hung KH, Moore SJ et al (2009) Paraphyly of organelle DNAs in *Cycas* Sect.
464 *Asiorientales* due to ancient ancestral polymorphisms. BMC Evol Biol 9:161
465 Doyle J (1991) DNA protocols for plants—CTAB total DNA isolation. In: Hewitt GM, Johnston
466 A (eds) Molecular techniques in taxonomy. Springer, Berlin
467 Drummond AJ, Suchard MA, Xie D et al (2012) Bayesian phylogenetics with BEAUti and the
468 BEAST 1.7. Mol Biol Evol 29:1969-1973
469 Duan L, Ming QZ, Zhang H et al (2013) Research progress of landslide dam events of Jinsha
470 River and its geomorphologic and environmental effects. Adv Geosci 3:8-17
471 Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for
472 visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour
473 4:359-361
474 Etterson JR (2004) Evolutionary potential of *Chamaecrista fasciculata* in relation to climate
475 change. II. Genetic architecture of three populations reciprocally planted along an environmental
476 gradient in the great plains. Evolution 58:1459-1471
477 Fjeldsa J, Lovett J (1997) Geographical patterns of old and young species in African forest biota:
478 The signifcance of specifc montane areas as evolutonary centres. Biodivers Conserv 6:325-346
479 Forster M, Forster P, Watson, J (2007) NETWORK (version 4.2. 0.1): a software for population
480 genetics data analysis. Fluxus Technology Ltd, Clare Available online at: fluxus-
481 engineering.com
482 Frankham R, Bradshaw CJA, Brook BW (2014) Genetics in conservation management: Revised
483 recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biol
484 Conserv 170:56-63
485 Fredrik R, Maxim T, Paul VDM et al (2012) MrBayes 3.2: Efficient Bayesian phylogenetic
486 inference and model choice across a large model space. Syst Biol 61:539-542
487 Gong X, Luan SS, Hung KH et al (2011) Population structure of *Nouelia insignis* (Asteraceae),
488 an endangered species in southwestern China, based on chloroplast DNA sequences: recent
489 demographic shrinking. J Plant Res 124:221-230
490 Gong YQ, Gong X (2016) Pollen-mediated gene flow promotes low nuclear genetic
491 differentiation among populations of *Cycas debaoensis* (Cycadaceae). Tree Genet Genomes
492 12:93

493 Graur D, Li WH (2000) Fundamentals of molecular evolution. Sinauer Associates, Sunderland.
494 Griffith MP, Calonje M, Meerow AW et al (2017) Will the same ex situ protocols give similar
495 results for closely related species? Biodivers Conserv 26:2951-2966
496 Griffith MP, Calonje M, Meerow AW et al (2015) Can a botanic garden cycad collection capture
497 the genetic diversity in a wild population? Int J Plant Sci 176:1-10
498 Hall JA, Walter GH (2011) Does pollen aerodynamics correlate with pollination vector? Pollen
499 settling velocity as a test for wind versus insect pollination among cycads (Gymnospermae:
500 Cycadaceae: Zamiaceae). Biol J Linn Soc 104:75-92
501 Hamrick JL, Godt MJW (1996) Conservation genetics of endemic plant species. In:Rieseberg
502 LH, Swensen SM (eds) Conservation genetics case histories from nature, Chapman & Hall New
503 York, 281-304.
504 Han J, Wang MC, Qiu Q et al (2016) Characterization of the complete chloroplast genome of
505 *Cycas panzhihuaensis*. Conserv Genet Resour 9:21-23
506 Hao YQ (2011) Florogeographical Analysis of Spermatophytes in *Cycas panzhihuaensis*
507 Assemblage. J Sichuan Forest Sci Technol 32:29-33
508 He YH, Li ZL (1999) The Ecological Geographic Distribution, Spatial Pattern And Collecting
509 History of *Cycas panzhihuaensis* Populations. Acta Phytoecol Sin 23:23-30
510 Hill K, Chen C, Loc P (2003a) Chapter 5: regional overview: Asia. In:Donaldson J (ed) Cycads:
511 status survey and conservation action plan. IUCN/SSC Cycad Specialist Group, IUCN, Gland,
512 Switzerland and Cambridge, pp 25–30.
513 Hill KD, Chase MW, Stevenson DW et al (2003b) The families and genera of cycads: A
514 molecular phylogenetic analysis of cycadophyta based on nuclear and plastid DNA sequences.
515 Int J Plant Sci 164:933-948
516 Huang S, Chiang YC, Schaal BA et al (2001) Organelle DNA phylogeography of *Cycas*
517 *taitungensis*, a relict species in Taiwan. Mol Ecol 10:2669-2681
518 IUCN (2017) The IUCN Red List of Threatened Species. Version 2017-3.
519 http://wwwiucnredlistorg Downloaded on 05 December 2017
520 James HE, Forster PI, Lamont RW et al (2018) Conservation genetics and demographic analysis
521 of the endangered cycad species *Cycas megacarpa* and the impacts of past habitat fragmentation.
522 Aust J Bot 66:173-189
523 Jetz W, Rahbek C, Colwell R (2004) The coincidence of rarity and richness and the potential
524 signature of history in centres of endemism. Ecol Lett 7:1180-1191
525 Jia J, Wu H, Wang JF et al (2014) Genetic diversity and structure of *Munronia delavayi* Franch.
526 (Meliaceae), an endemic species in the dry-hot valley of Jinsha River, south-western China.
527 GENET RESOUR CROP EV 61:1381-1395
528 Jump AS, Penuelas J (2005) Running to stand still: adaptation and the response of plants to rapid
529 climate change. Ecol Lett 8:1010-1020
530 Kark S, Alkon PU, Safriel U et al (1999) Conservation priorities for chukar partridge in israel
531 based on genetic diversity across an ecological gradient. Conserv Biol 13:542-552

532 Kaulfuss F, Reisch C (2017) Reintroduction of the endangered and endemic plant species
533 *Cochlearia bavarica*-implications from conservation genetics. Ecol Evol 7:11100-11112
534 Kovach W (1998) MVSP-A multivariate statistical Package for Windows, ver. 3.1. Kovach
535 Comput Serv, Wales, UK 137.
536 Lande R (1999) Genetics and the extinction of species. In:Landweber LF, Dobson AP (eds)
537 Extinction risks from anthropogenic, ecological and genetic factors, Princeton University Press
538 edn.
539 Larson WA, Seeb LW, Everett MV et al (2014) Genotyping by sequencing resolves shallow
540 population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*).
541 Evol Appl 7:355-369
542 Li CL, Wang Q, Jiang SY et al (1999) Genetic diversity of allozymes in populations of *Cycas*
543 *panzhihuaensis* L. Zhou & S.Y. Yang. In: *Proceedings of the 9th International Conference on*
544 *Cycad Biology*, pp 323-327.
545 Li L, Wang ZF, Jian SG et al (2009) Isolation and characterization of microsatellite loci in
546 endangered *Cycas changjiangensis* (Cycadaceae). Conserv Genet 10:793-795
547 Li YY, Tsang EPK, Cui MY et al (2012) Too early to call it success: An evaluation of the natural
548 regeneration of the endangered *Metasequoia glyptostroboides*. Biol Conserv 150:1-4
549 Liu GF, Deng TX, Zhou L et al (1990) Karyotype Analysis of *Cycas panzhihuaensis* L. Zhou &
550 S.Y.Yang. Chin J Plant Ecol 14:23-30
551 Liu H, Lin WY, Li XY (2016) Chinese nature reserve continues experimental use of fire to
552 benefit the threatened *Cycas panzhihuaensis*. Oryx 50:582
553 Liu J (2016) Molecular Phylogeny and Biogeography of *Cycas* L. Kunming Institute of Botany,
554 Chinese Academy of Sciences
555 Liu J, Zhang SZ, Nagalingum NS et al (2018) Phylogeny of the gymnosperm genus *Cycas* L.
556 (Cycadaceae) as inferred from plastid and nuclear loci based on a large-scale sampling:
557 Evolutionary relationships and taxonomical implications. Mol Phylogen Evol 127:87-97
558 Liu J, Zhou W, Gong X (2015) Species delimitation, genetic diversity and population historical
559 dynamics of *Cycas diannanensis* (Cycadaceae) occurring sympatrically in the Red River region
560 of China. Frontiers Plant Sci 6:696
561 Mankga LT, Yessoufou K (2017) Factors driving the global decline of cycad diversity. AoB
562 Plants 9:plx022
563 Mouillot D, Parravicini V, Bellwood DR et al (2016) Global marine protected areas do not
564 secure the evolutionary history of tropical corals and fishes. Nat Commun 7:10359
565 Myers N, Mittermeier RA, Mittermeier CG et al (2000) Biodiversity hotspots for conservation
566 priorities. Nature 403:853-858
567 Nagalingum NS, Marshall CR, Quental TB et al (2011) Recent synchronous radiation of a living
568 fossil. Science 334:796
569 Nash HC, Wirdateti, Low GW et al (2018) Conservation genomics reveals possible illegal trade
570 routes and admixture across pangolin lineages in Southeast Asia. Conserv Genet
571 Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, USA.

572 Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic
573 software for teaching and research. Mol Ecol Resour 6:288-295
574 Pérez-Espona S (2017) Conservation genetics in the European Union-Biases, gaps and future
575 directions. Biol Conserv 209:130-136
576 Petit R, Abdelhamid EM, Pons O (2008) Identifying populations for conservation on the basis of
577 genetic markers. Conserv Biol 12:844-855
578 Pierson JC, Coates DJ, Oostermeijer JGB et al (2016) Genetic factors in threatened species
579 recovery plans on three continents. Front Ecol Environ 14:433-440
580 Piry S, Luikart G, Cornuet J (1999) Computer note. BOTTLENECK: a computer program for
581 detecting recent reductions in the effective size using allele frequency data. Hereditas 90:502-
582 503
583 Pons, Peti RJ (1996) Measwring and testing genetic differentiation with ordered versus
584 unordered alleles. Genetics 144:1237-1245
585 Pritchard JK, Stephens M, Rosenberg NA et al (2000) Association mapping in structured
586 populations. Am J Hum Genet 67:170
587 Rambaut A, Suchard M, Drummond A (2014) Tracer v1. 6. chromosomal evolution in higher
588 plants. Available online at: http://tree.bio.ed.ac.uk/software/tracer/
589 Rauch EM, Bar-Yam Y (2004) Theory predicts the uneven distribution of genetic diversity
590 within species. Nature 431:449-452
591 Reisch C, Bernhardt-Römermann M (2014) The impact of study design and life history traits on
592 genetic variation of plants determined with AFLPs. Plant Ecol 215:1493-1511
593 Reisch C, Schmidkonz S, Meier K et al (2017) Genetic diversity of calcareous grassland plant
594 species depends on historical landscape configuration. BMC Ecol 17:19
595 Rozas J (2009) DNA Sequence Polymorphism analysis using DnaSP. Methods Mol Biol
596 537:337-350
597 Schindler DE, Hilborn R, Chasco B et al (2010) Population diversity and the portfolio effect in
598 an exploited species. Nature 465:609
599 Schneider S, Roessli D, Excoffier L et al (2000) Arlequin: a software for population genetics
600 data analysis, Version 2.000. Lilog
601 Segarra-Moragues J, Palop-Esteban M, Gonzalez-Candelas F et al (2005) On the verge of
602 extinction: Genetics of the critically endangered Iberian plant species, *Borderea chouardii*
603 (Dioscoreaceae) and implications for conservation management. Mol Ecol 14:969-982
604 Selwood KE, McGeoch MA, Mac Nally R (2015) The effects of climate change and land-use
605 change on demographic rates and population viability. Biol Rev Camb Philos Soc 90:837-853
606 Shaw J, Lickey E, Beck JT (2005) The tortoise and the hare II: Relative utility of 21 noncoding
607 chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92:142-166
608 Swindell SR, Plasterer TN (1997) SEQMAN. Springer New York
609 Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (and other methods),
610 Version 4.0b10. Sinauer Associates, Inc, Sunderland, Massachusetts.

611 Tang YL, Su ZX (2004) The research actuality and prospect of a rare and endangered plant
612 named *Cycas panzhihuaensis*. J Shaanxi Normal Univ (Nat Sci Ed) 18:87-92
613 Trifinopoulos J, Nguyen LT, Haeseler AV et al (2016) W-IQ-TREE: a fast online phylogenetic
614 tool for maximum likelihood analysis. Nucleic Acids Res 44:W232-W235
615 Wang HC (2016) Characteristics of *Cycas Panzhihuaensis* Community in Pudu Valley of
616 Yunnan. Forest Inv Plan 41:32-39
617 Wang J, Wang YY, Zhou H et al (2010) ISSR analysis on genetic diversity of *Cycas*
618 *panzhihuaensis* L. Zhou et S. Y. Yang. J Sichuan Univ (Nat Sci Ed) 47:366-370
619 Wang Q, Li CL, Yang SY et al (1997) Pollination biology of *Cycas panzhihuaensis* L. Zhou et S.
620 Y. Yang. Acta Bot Sin 39:156-163
621 Wang YY, Zhou H, Su B et al (2011) Study on seed coat structure and trachary elements in
622 xylems of *Cycas panzhihuaensis* L. Zhou et S. Y. Yang. J Sichuan Univ (Nat Sci Ed) 48:247-252
623 Wang ZF, Ye WH, Cao HL et al (2008) Identification and characterization of EST-SSRs and
624 cpSSRs in endangered *Cycas hainanensis*. Conserv Genet 9:1079-1081
625 Waples RS, Do C (2008) ldne: a program for estimating effective population size from data on
626 linkage disequilibrium. Mol Ecol Resour 8:753-756
627 Wright S (1931) Evolution in Mendelian populations. Genetics 16:97
628 Yang R, Feng X, Gong X (2017) Genetic structure and demographic history of *Cycas chenii*
629 (Cycadaceae), an endangered species with extremely small populations. Plant Diver 39:44-51
630 Yang Y, Li Y, Li LF et al (2008) Isolation and characterization of microsatellite markers for
631 *Cycas debaoensis* Y. C. Zhong et C. J. Chen (Cycadaceae). Molecular Ecology Resources 8:913-
632 915
633 Yang YQ, Huang BH, Yu ZX et al (2015) Inferences of demographic history and fine-scale
634 landscape genetics in *Cycas panzhihuaensis* and implications for its conservation. Tree Genet
635 Genom 11:78
636 Yessoufou K, Daru BH, Tafirei R et al (2017) Integrating biogeography, threat and evolutionary
637 data to explore extinction crisis in the taxonomic group of cycads. Ecol Evol 7:2735-2746
638 Yu WB (2015) Floral and reproductive ecology of *Cycas panzhihuaensis*. ATBC Asia Pacific
639 Chapter Meeting
640 Yu ZX, Yang YQ, Liu J et al (2007) Preliminary experimenton propagation of *Cycas*
641 *panzhihuaensis*. J Southwest Forestry Coll 27:36-41
642 Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on greenhouse
643 warming and carbon-cycle dynamics. Nature 451:279-283
644 Zhang F, Su T, Yang Y et al (2010) Development of seven novel EST-SSR markers from *Cycas*
645 *panzhihuaensis* (Cycadaceae). Am J Bot 97:e159-161
646 Zhang FM (2012) Conservation Genetics of *Cycas panzhihuaensis*. Kunming Institute of Botany,
647 Chinese Academy of Sciences
648 Zhang Q, Zhao Y, Gong X (2011a) Genetic variation and phylogeography of *Psammosilene*
649 *tunicoides* (Caryophyllaceae), a narrowly distributed and endemic species in south-western
650 China. Aust J Bot 59:450-459

651 Zhang TC, Comes HP, Sun H (2011b) Chloroplast phylogeography of *Terminalia franchetii*
652 (Combretaceae) from the eastern Sino-Himalayan region and its correlation with historical river
653 capture events. Mol Phylogenet Evol 60:1-12
654 Zhao YJ, Gong X (2015) Diversity and conservation of plant species in dry valleys, southwest
655 China. Biodivers Conserv 24:2611-2623
656 Zheng Y, Liu J, Gong X (2016) Tectonic and climatic impacts on the biota within the Red River
657 Fault, evidence from phylogeography of *Cycas dolichophylla* (Cycadaceae). Sci Rep 6:33540
658 Zhou H, Wang J, Huang R et al (2009) Callus induction and browning inhibition in *Cycas*
659 *panzhihuaensis* L. Zhou et S. Y. Yang. J Sichuan Univ (Nat Sci Ed) 46:849-852
660 Zhou L, Yang SY, Fu LG et al (1981) Two new species of *Cycas* from Sichuan. Acta Phytotaxon
661 Sin 19:335-338

Table 1(on next page)

Parameters of genetic diversity and genetic differentiation for the combined cpDNA, nDNA *PHYP* and combined three nDNA markers (N3) in seven populations of *C. panzhihuaensis*

Hd: Haplotype (gene) diversity; $\boldsymbol{\pi}$: Nucleotide diversity; $\boldsymbol{H_S}$: Sequence diversity within population; $\boldsymbol{H_T}$: Total sequence diversity; $\boldsymbol{G_{ST}}$, $\boldsymbol{N_{ST}}$: The coefficient of genetic differentiation for non-ordered alleles.

Marker	*Hd*	π $\times 10^{-3}$	H_S	H_T	G_{ST}	N_{ST}	U test
cpDNA	0.8033	2.59	0.235 (0.1168)	0.881 (0.0926)	0.733 (0.1314)	0.782 (0.1298)	non-significant
PHYP	0.3302	0.4	0.257 (0.1025)	0.366 (0.1133)	0.296 (0.0480)	0.296 (0.0652)	non-significant
N3	0.7350	0.56	0.506 (0.1016)	0.769 (0.1078)	0.342 (0.0919)	0.420 (0.1202)	non-significant

1

Table 2 (on next page)

Parameters of genetic diversity and genetic differentiation for 11 microsatellite markers in seven populations of C. *panzhihuaensis*

N: sample size, ***Na***: number of alleles, ***Ne***: number of effective alleles, ***I***: Shannon's index, ***Ho***: observed heterozygosity, ***He***: expected heterozygosity, ***F***: fixation index, ***%P***: percentage of polymorphic loci. *N*e of population JPD and GH could not be detected because of the insufficient sample size.

Pop	*N*	*%P*	*Na*		*Ne*		*I*		*Ho*		*He*		*F*		*Ne*
			Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	
DSS	20	63.64%	1.909	0.251	1.605	0.179	0.461	0.116	0.286	0.095	0.298	0.075	0.061	0.160	0.9
PDH	20	72.73%	2.545	0.455	1.589	0.204	0.515	0.137	0.118	0.028	0.282	0.073	0.456	0.100	3.3
PZH	20	72.73%	3.273	0.662	1.861	0.409	0.606	0.175	0.182	0.060	0.306	0.082	0.440	0.126	17.6
WQ	20	81.82%	2.818	0.536	1.953	0.379	0.630	0.162	0.236	0.064	0.353	0.081	0.347	0.117	15.4
SL	11	90.91%	3.273	0.524	2.373	0.374	0.825	0.179	0.107	0.042	0.445	0.090	0.643	0.126	1.8
JPD	6	72.73%	2.091	0.315	1.835	0.265	0.552	0.139	0.227	0.079	0.347	0.081	0.246	0.180	-
GH	5	54.55%	1.909	0.315	1.541	0.178	0.428	0.134	0.164	0.065	0.262	0.079	0.302	0.165	-
Total	102	72.73%	2.545	0.178	1.822	0.113	0.574	0.056	0.189	0.025	0.328	0.030	0.374	0.054	39.0

1

Table 3(on next page)

Results of analysis of molecular variance (AMOVA) of combined cpDNA, nuclear gene *PHYP*, combined nDNA sequences (N3)sequences and microsatellite markers of *C. panzhihuaensis.*

d.f.: degrees of freedom; $\boldsymbol{F_{ST}}$: fixation index; ***Nm***: gene flow = $(1-F_{ST})/4*F_{ST}$

Markers	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation (%)	F_{ST}	*N*m
cpDNA	Among populations	6	157.166	2.980	86.60	0.866	0.039
	Within populations	54	24.900	0.461	13.40		
PHYP	Among populations	6	211.481	1.824	32.73	0.327	0.514
	Within populations	115	431.150	3.749	67.27		
N3	Among populations	6	29.848	0.266	40.83	0.408	0.362
	Within populations	115	44.300	0.385	59.17		
microsatellite	Among populations	6	165.258	0.912	33.06	0.331	0.506
	Within populations	197	363.830	1.847	66.94		

1

Table 4(on next page)

Parameters of neutrality tests and mismatch analysis based on cpDNA and nuclear genes of *C. panzhihuaensis*.

SSD: the sum-of-squared deviations; **R**: raggedness index; **#**: $P < 0.05$ significant difference.

PeerJ Preprints NOT PEER-REVIEWED

Marker	**Tajima's** D	**Fu and Li's** D*	**Fu and Li's** F*	**Fu's** Fs	SSD	R
cpDNA	0.493325	1.82738	1.59999	13.648	0.06186	0.16882
PHYP	-1.19639	-0.06795	-0.52146	-2.021	0.20020#	0.34830
N3	-0.19000	1.17849	0.84329	-2.791	0.00920	0.11837

1

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27265v1 | CC BY 4.0 Open Access | rec: 8 Oct 2018, publ: 8 Oct 2018

PeerJ Preprints NOT PEER-REVIEWED

Table 5(on next page)

Bottleneck analysis and Garza-Williamson index based on microsatellite loci from six populations of C. *panzhihuaensis*

*: P< 0.05, significant difference; **: P< 0.01, highly significant difference. Indices of population GH could not be detected by Bottleneck analysis because of its insufficient sample size.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27265v1 | CC BY 4.0 Open Access | rec: 8 Oct 2018, publ: 8 Oct 2018

PeerJ Preprints NOT PEER-REVIEWED

Population	T.P.M		S.M.M		Mode shift	Garza-Williamson index
	Sign test	Wilcoxon test	Sign test	Wilcoxon test		
DSS	0.01328*	0.00073**	0.01422*	0.02100*	shifted mode	0.45034
PDH	0.08215*	0.17480	0.43846	0.51953	L-shaped distribution	0.41138
PZH	0.34274	0.57715	0.33621	0.70020	L-shaped distribution	0.43259
WQ	0.22812	0.06738*	0.47564	0.36523	L-shaped distribution	0.39331
SL	0.02812*	0.00098**	0.02932*	0.00684**	L-shaped distribution	0.45275
JPD	0.02284*	0.00342**	0.03018*	0.00684**	shifted mode	0.39471
GH	-	-	-	-	-	0.34077
Mean	-	-	-	-	-	0.41083

PeerJ Preprints NOT PEER-REVIEWED

Figure 1

Geographical distributions of haplotypes identified in seven populations of *C. panzhihuaensis*.

(a) cpDNA; (b) nDNA *PHYP*; (c) N3. Blue line represents Jinsha River. Frequencies of haplotypes in each population are indicated by the proportions of pie diagrams. Maps were drawn by ArcGIS v10.2 (http:// desktop.arcgis.com).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27265v1 | CC BY 4.0 Open Access | rec: 8 Oct 2018, publ: 8 Oct 2018

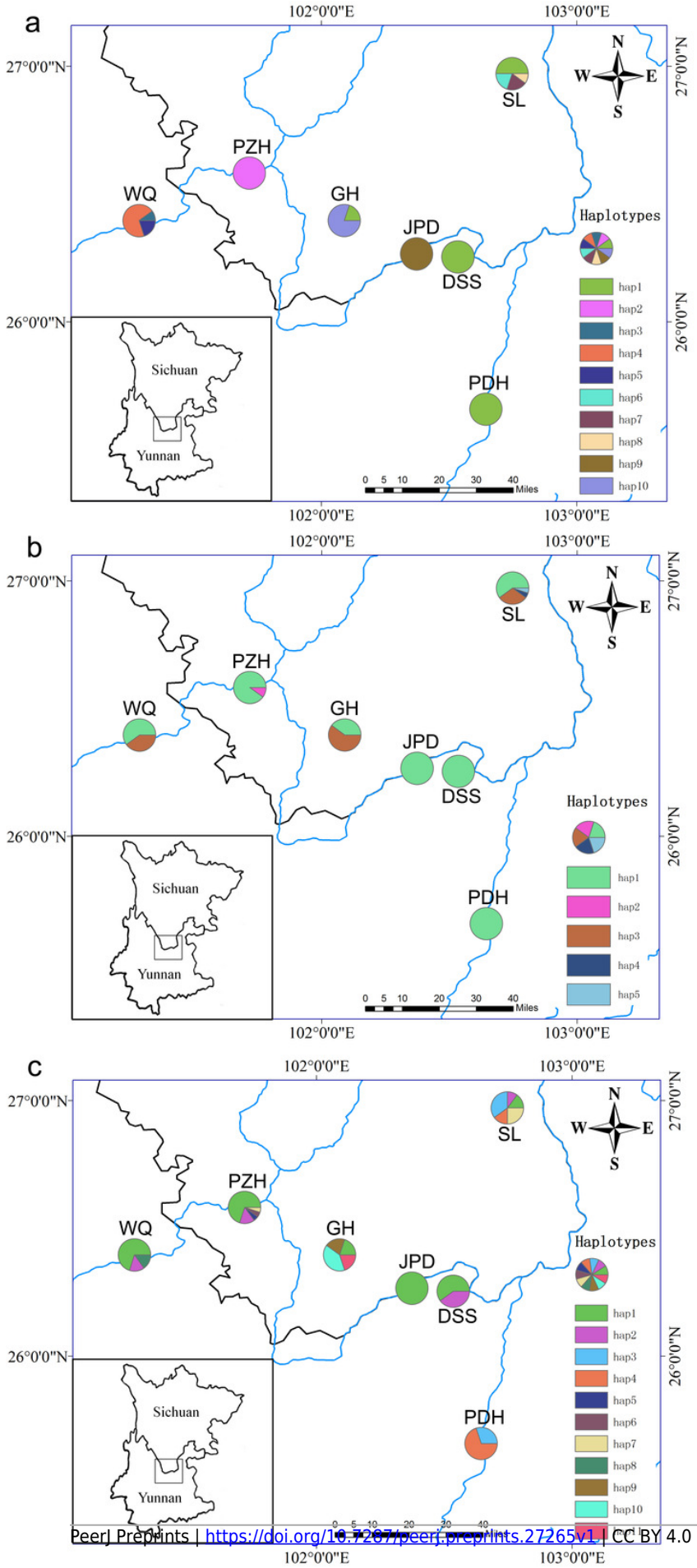
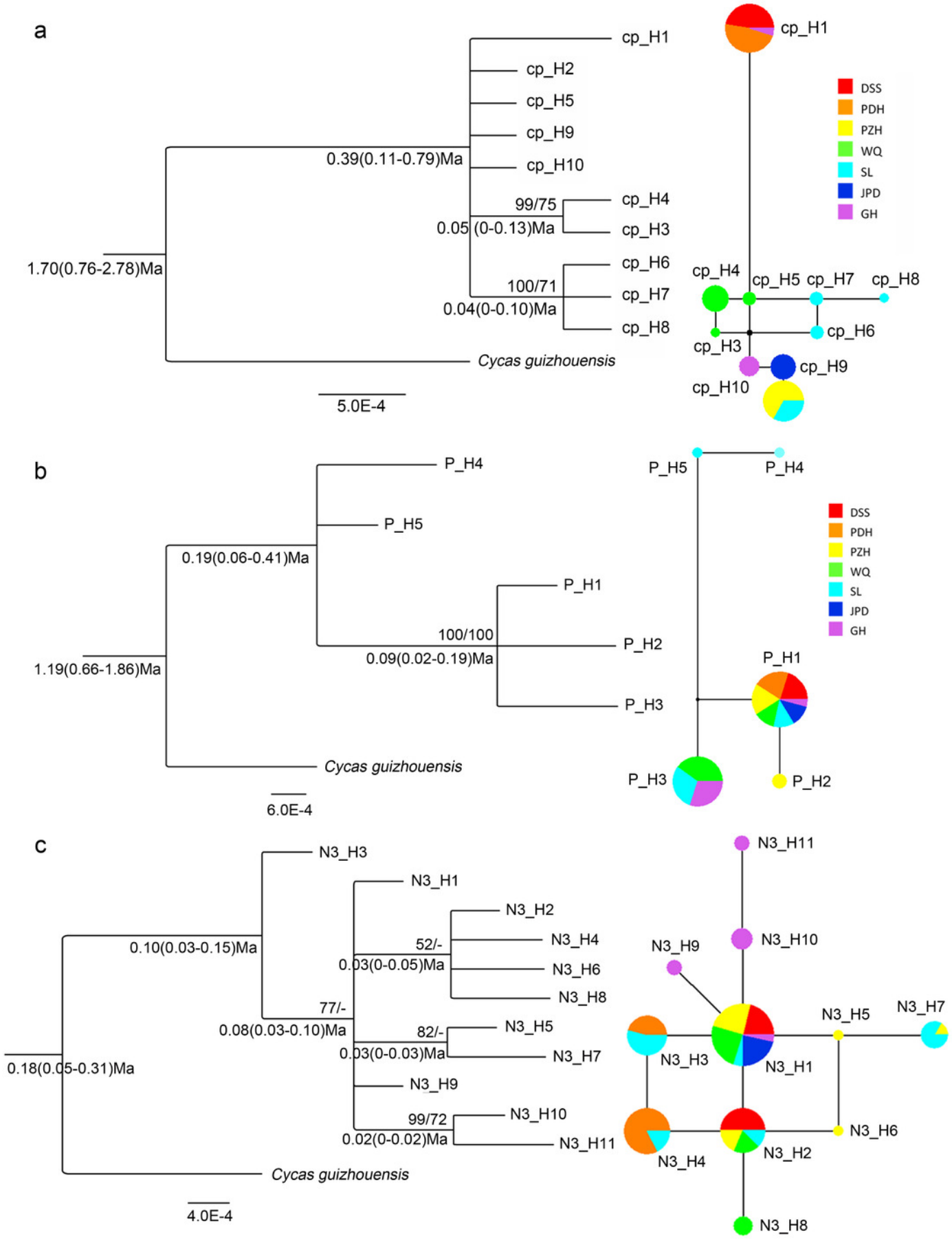

a
102°0'0"E
103°0'0"E
27°0'0"N
26°0'0"N
N
W
E
S
SL
PZH
WQ
GH
JPD
DSS
PDH
Haplotypes
hap1
hap2
hap3
hap4
hap5
hap6
hap7
hap8
hap9
hap10
Sichuan
Yunnan
0 5 10 20 30 40
Miles
b
hap1
hap2
hap3
hap4
hap5
c
hap1
hap2
hap3
hap4
hap5
hap6
hap7
hap8
hap9
hap10

Figure 2

Phylogenetic relationship among haplotypes and median-joining network of haplotypes.

(Left) Phylogenetic relationship among haplotypes based on Bayesian Inference (BI) and Maximum Parsimony. Method based on different datasets, with *Cycas guizhouensis* used as outgroup. Numbers on branches are the posterior probability (PP) and bootstrap values (BP) respectively, numbers below branches indicate the divergence time with 95% highest posterior density (HPD). (Rignt) Median-joining network of haplotypes inferred from different datasets. Color and size of pie correspond to populations and proportion of individuals respectively. (a) cpDNA; (b) *PHYP*; (c) N3.


a
cp_H1
cp_H2
cp_H5
cp_H9
cp_H10
0.39(0.11-0.79)Ma
99/75
cp_H4
cp_H3
0.05 (0-0.13)Ma
cp_H6
100/71
cp_H7
0.04(0-0.10)Ma
cp_H8
1.70(0.76-2.78)Ma
Cycas guizhouensis
5.0E-4
DSS
PDH
PZH
WQ
SL
JPD
GH
b
P_H4
P_H5
0.19(0.06-0.41)Ma
P_H1
100/100
P_H2
0.09(0.02-0.19)Ma
P_H3
1.19(0.66-1.86)Ma
Cycas guizhouensis
6.0E-4
c
N3_H3
N3_H1
N3_H2
52/-
N3_H4
0.03(0-0.05)Ma
N3_H6
N3_H8
0.10(0.03-0.15)Ma
77/-
0.08(0.03-0.10)Ma
82/-
N3_H5
0.03(0-0.03)Ma
N3_H7
0.18(0.05-0.31)Ma
N3_H9
99/72
N3_H10
0.02(0-0.02)Ma
N3_H11
Cycas guizhouensis
4.0E-4
N3_H11
N3_H10
N3_H9
N3_H5
N3_H7
N3_H3
N3_H1
N3_H6
N3_H4
N3_H2
N3_H8

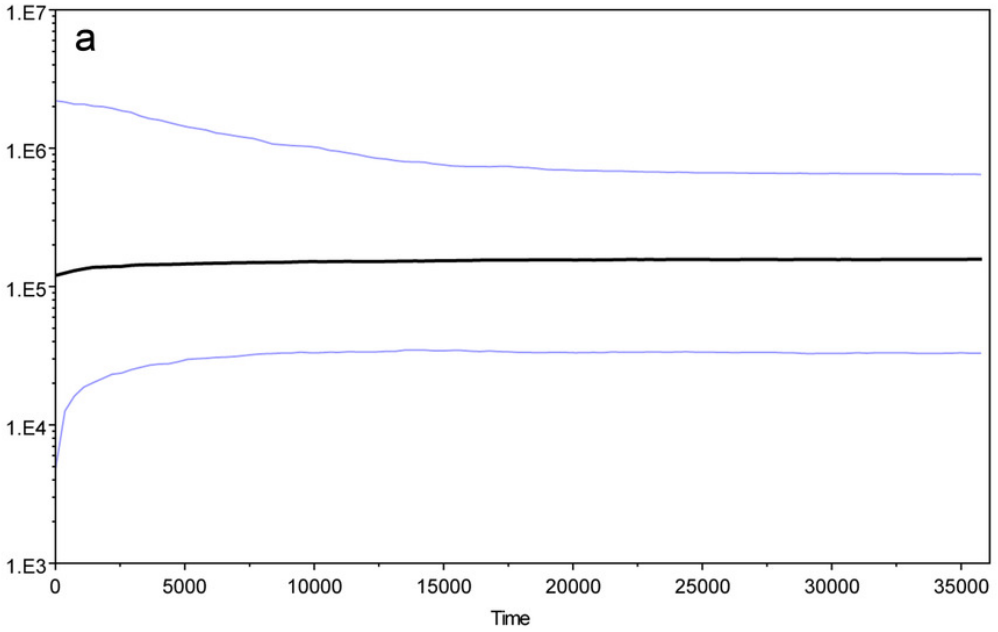
PeerJ Preprints NOT PEER-REVIEWED

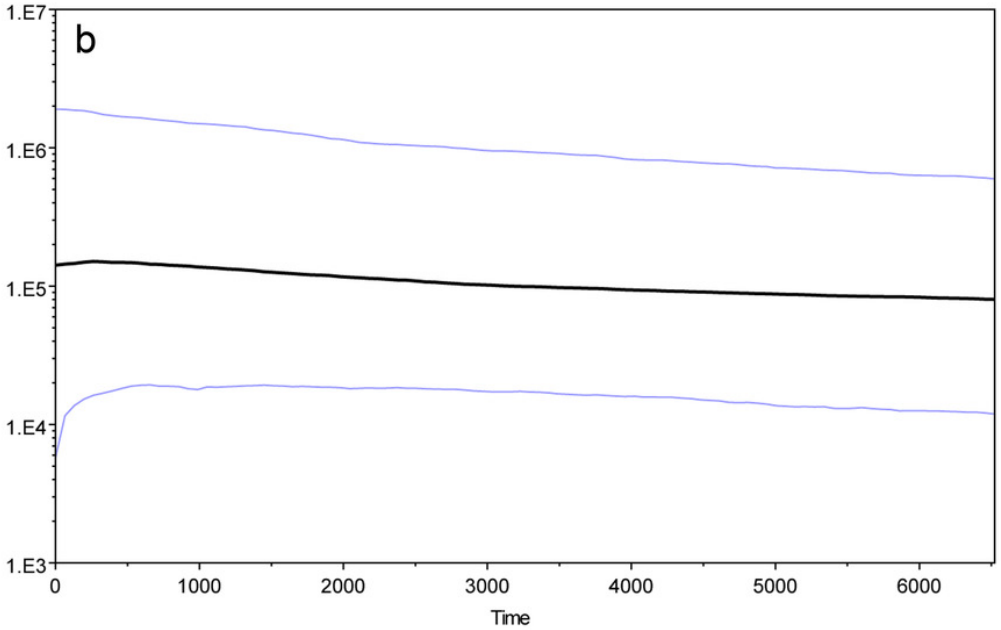
Figure 3

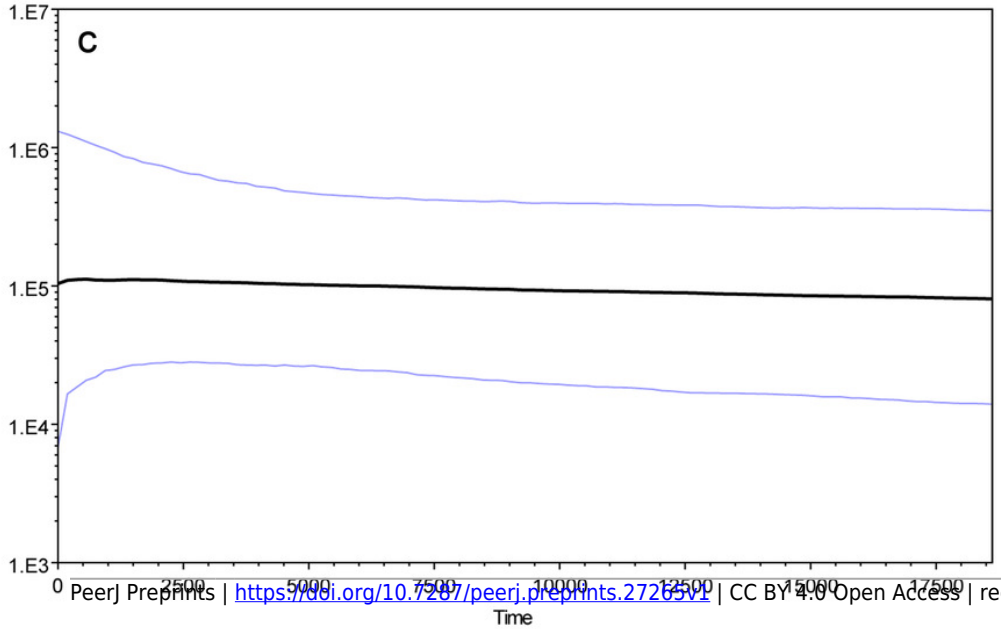
Bayesian Skyline Plots based on cpDNA and nuclear genes.

(a) cpDNA; (b) *PHYP*; (c) N3. The plots reflect the effective population size fluctuation throughout time. Black line: median estimation; area between blue lines: 95% confidence interval.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27265v1 | CC BY 4.0 Open Access | rec: 8 Oct 2018, publ: 8 Oct 2018


a
1.E7
1.E6
1.E5
1.E4
1.E3
0
5000
10000
15000
20000
25000
30000
35000
Time


b
1.E7
1.E6
1.E5
1.E4
1.E3
0
1000
2000
3000
4000
5000
6000
Time


c
1.E7
1.E6
1.E5
1.E4
1.E3
0
2500
5000
7500
10000
12500
15000
17500
Time