How much genetic variation is stored in the endangered and fragmented shrub *Tetraena mongolica* Maxim?

Yingbiao Zhi Corresp., 1, 2, 3 , **Zhonglou Sun** 2 , **Ping Sun** 2 , **Kai Zhao** 4 , **Yangnan Guo** 5, 6 , **Dejian Zhang** 5 , **Baowei Zhang** Corresp. 2

1 School of Ecology and Environment, Inner Mongolia University, Hohhot, Inner Mongolia, China
2 School of Life Sciences, Anhui University, Hefei, Anhui, China
3 Ordos Institute of Technology, Ordos, Inner Mongolia, China
4 School of resource and environment, Anqing Normal University, Anqing, Anhui, China
5 School of Life Sciences, Inner Mongolia University, Hohhot, Inner Mongolia, China
6 China Energy Technology Holdings Ltd., Beijing, China

Corresponding Authors: Yingbiao Zhi, Baowei Zhang
Email address: zyb.china@163.com, zhangbw@ahu.edu.cn

Tetraena mongolica Maxim (Zygophyllaceae) is an endangered species endemic to western Inner Mongolia and China, and is currently threatened by habitat loss and human over-exploitation. We explored the genetic background, its genetic diversity, population structure, and demographic history, based on 12 polymorphic nuclear microsatellite loci. Our results indicated high genetic diversity in extant populations, but no distinguishable gene cluster corresponding with a specific biogeography. Population demography analysis using a MSVAR indicated a strong, recent population decline approximately 5,455 years ago. These results suggest that the Yellow River and Zhuozi Mountain range may not prevent pollination between populations. Finally, we surmised that the population demography of *T. mongolica* was likely to have been affected by early mankind activities.

1 **How Much Genetic Variation is Stored in the Endangered and**

2 **Fragmented Shrub *Tetraena mongolica* Maxim?**

4 Yingbiao Zhi[1,2,3,*,†], Zhonglou Sun[2,*], Ping Sun[2], Kai Zhao[4], Yangnan Guo[5,6], Dejian Zhang[5],

5 Baowei Zhang[2,†]

7 1. School of Ecology and Environment, Inner Mongolia University, Hohhot, Inner Mongolia,

8 China

9 2. School of Life Sciences, Anhui University, Hefei, Anhui, China

10 3. Ordos Institute of Technology, Ordos, Inner Mongolia, China

11 4. School of resource and environment, Anqing Normal University, Anqing, Anhui, China

12 5. School of Life Sciences, Inner Mongolia University, Hohhot, Inner Mongolia, China

13 6. China Energy Technology Holdings Ltd., Beijing, China

15 * These authors contributed equally to this work.

17 Corresponding author:

18 Yingbiao Zhi,

19 Address: 235 West University Road, Hohhot, Inner Mongolia, 010021, China

20 E-mail: zyb.china@163.com

21 Baowei Zhang,

22   Address: 111 Jiulong Road, Hefei, Anhui, 230601, China,

23   E-mail: zhangbw@ahu.edu.com

24 **ABSTRACT**

25   *Tetraena mongolica* Maxim (Zygophyllaceae) is an endangered species endemic to western

26   Inner Mongolia and China, and is currently threatened by habitat loss and human over-

27   exploitation. We explored the genetic background, its genetic diversity, population structure, and

28   demographic history, based on 12 polymorphic nuclear microsatellite loci. Our results indicated

29   high genetic diversity in extant populations, but no distinguishable gene cluster corresponding

30   with a specific biogeography. Population demography analysis using a MSVAR indicated a

31   strong, recent population decline approximately 5,455 years ago. These results suggest that the

32   Yellow River and Zhuozi Mountain range may not prevent pollination between populations.

33   Finally, we surmised that the population demography of *T. mongolica* was likely to have been

34   affected by early mankind activities.

INTRODUCTION

36   Understanding population history and genetic structure is a key aspect of ecological

37   research (Rockwood, 2006). Endemic species with restricted geographic distributions have

38   become a central concern of biologists faced with the problem of preserving rare species

39   endangered by habitat destruction and fragmentation (Ge et al., 2003). For endemics with narrow

40   ranges and declining populations, information about historical patterns of demography, genetic

41   structure, and genetic variation within and among natural populations helps to clarify population

42   structure, and the organism's evolutionary history, supporting conservation and management

43   efforts (Moritz, 1994; Ge et al., 2011). Intraspecific genetic variation is the most fundamental

44   level of biodiversity, providing the basis for evolutionary change and the ability of species to

45   adapt to new environmental conditions (Frankham, Ballou & Briscoe, 2002). In contrast, plants

46   and invertebrates with high fecundity and human-mediated dispersal ability, its populations can

47   be successfully reestablished and will experience range expansion (Song et al., 2013).

48   Previous studies revealed that natural landscape features such as mountains and rivers can

49   act as genetic boundaries and shape the structure of populations (Funk et al., 2001; Whiteley et

50   al., 2004). However, anthropogenic landscape features also have an impact on genetic structure

51   (Gaines et al., 1997) and population dynamics (Nupp & Swhart, 1998). Anthropogenic

52   disturbance, such as roads, have dramatically increased the physical isolation of populations, and

53   it has been assumed that such isolation will lead to reduced gene flow and consequently reduced

54   genetic diversity in populations (Byrne et al., 2007). Furthermore, habitat destruction and land-

55   use change may also influence gene flow at the landscape scale (Manel et al., 2003; Eckstein et

56 al., 2006). These anthropogenic effects may also occur in the center of a species' range and may

57 thus be superimposed on natural geographic patterns.

58   *Tetraena mongolica* Maxim is a member of the broader genus *Tetraena* in the subfamily

59 Zygophyllaceae (Beier et al., 2003; Lauterbach et al., 2016), and is endemic to western Inner

60 Mongolia around the Yellow River basin, and is nationally endangered in China (Fu, 1992; Xu et

61 al., 1998; Zhang & Yang, 2000). Its distribution is restricted to the western Gobi, the largest

62 desert in Asia, and one characterized by extremely low annual rainfall (Xu et al., 1998; Zhang &

63 Yang, 2000), where *T. mongolica* is able to survive because of its extensive root system, and acts

64 as a windbreak and soil stabilizer (Dong & Zhang, 2001; Zhang et al., 2003). Its stems contain

65 high levels of waxes and oils (Wang, Ma & Zheng, 2000), and are combustible, even when green.

66 For this reason, *T. mongolica* is a popular firewood species, and its range has declined

67 alarmingly through overexploitation (Zhang & Yang, 2000; Ge et al., 2011). Based on inter-

68 simple sequence repeats (ISSR) marker, Ge et al., (2003) revealed that this species presents an

69 intermediate level of intraspecific genetic diversity despite its limited distribution. Moreover, Ge

70 et al., (2003) discovered that there was low genetic differentiation among *T. Mongolic*

71 populations, which was due to the extensive gene flow within this population. However, neither

72 the impacts of natural barriers to dispersal, nor human influences on the genetic structure and

73 demographic history of *T. mongolica* have been ascertained.

74   Evolutionary, demographic and genetic analyses all contribute to conservation and

75 management of species (Beaumont & Bruford, 1999; O'Brien, 1994). We generated a

76 comprehensive genetic characterization for *T. mongolica* with the aim of supporting a

77 conservation strategy. We used twelve microsatellites SSRs (Simple Sequence Repeats)

78 genotyped onto an extensive dataset to evaluate the current genetic diversity in *T. mongolica*

79 populations, and to assess the effect of natural landscape barriers (Yellow River and Zhuozi

80 Mountains) in shaping population structure. Lastly, we modeled the demographic history of *T.*

81 *mongolica* to assess the effects of historic events on population demography. Our findings may

82 be useful for the conservation and management of *T. mongolica* and other species endemic to the

83 Yellow River basin.

84 **MATERIAL AND METHODS**

85 **Ethical statement**

86   The collection of samples was performed within an investigation project on plants of *T.*

87 *mongolica*. This investigation project and the sample collection were approved by the West

88 Ordos National Nature Reserve, Inner Mongolia Province, China. Field experiments were also

89 approved by the West Ordos National Nature Reserve, Inner Mongolia Province, China.

90 **Sample collection**

91   Between 2010 and 2014, 339 leaf samples of *T. mongolica* were collected from eight

92 populations along the G6 Road: Shizuishan (SZS, N = 32); Dishan (DS, N = 64); Hainan (HN, N

93 = 51); Dongalashan (DALS, N = 62); Wuda (WD, N = 32); Qianlishan (QLS, N = 32);

94 Wujiamiao (WJM, N = 36) and Taositu (TST, N = 30) (Fig. 1). Leaves sample were powdered in

95 liquid nitrogen and stored frozen at -80 °C.

96 **DNA extraction, PCR amplification and microsatellite genotyping**

97   Total genomic DNA was extracted from the powdered tissue following a modified CTAB

98 procedure (Doyle & Doyle, 1987), and purified via an EasyPure PCR Purification Kit

99 (TransGene). In the present study, we used twelve high polymorphic loci for *T. mongolica* (Zhi

100 et al., 2014) as genetic markers. PCR reaction mixtures (25 μL) consisted of 1 μL genomic DNA

101 (concentration 10 - 50 ng/μL), 2 μL 10× buffer, 1 μL of 2.5 mM MgSO$_4$, 2 μL of 2 mM dNTPs,

102 1 U *Taq* polymerase, 0.3 mM of each primer (forward primer fluorescently labeled with FAM,

103 HEX or TAMRA) and sufficient water. The amplification program was conducted with

104 following conditions: 5 min denaturing at 95°C; followed by 35 cycles of 30s at 95°C, 20 s at the

105 annealing temperature (55 / 60 °C), 30 s at 72°C; and 5 min at 72°C. PCR products were

106 genotyped on an ABI 3730 semi-automated sequencer (PE Applied Biosystems) utilizing the

107 GS500 marker, followed by analysis under GeneMarker 1.85 (SoftGenetics LLC) (Holland and

108 Parson, 2011).

109 **Data analysis**

110 The presence of null alleles and genotyping errors in microsatellite genotyping was detected

111 by Micro-Checker v2.2.3 as previously described by Van Oosterhout et al., (2004), while the

112 linkage disequilibrium was tested with GENEPOP 4.2.1 as described by Rousset (2008). In

113 addition, several population genetic summary statistics to describe genetic variation were

114 estimated by GENETIX v.4.02 as described in Belkhir, Borsa, Chikhi, Raufaste, & Bonhomme

115 (2001), including mean number of alleles per locus (MNA), observed heterozygosities (H_O),

116 expected heterozygosities (H_E) and inbreeding coefficients (F_{IS}). In addition, allelic richness (AR)

117 was also calculated to estimate the allelic diversity that compensates for unequal sample size by

118 FSTAT and averaged across loci (Goudet, 2001). Genetic differentiation (F_{ST}) between

119   populations was estimated using ARLEQUIN 3.0 (Excoffier, Laval, & Schneider, 2005), and

120   statistical significance of F_{ST} values was tested with 10,000 permutations. In addition, the

121   association between the estimates of $F_{ST} / 1 - F_{ST}$ (Rousset & Raymond, 1997) and land-based

122   Manhattan distance were assessed using the Mantel test, implemented in the Isolation by

123   Distance Web Service (IBDWS) software (Jensen, Bohonak, & Kelley, 2005); the statistical

124   significance of the values was obtained by 10,000 randomization steps.

125   A Bayesian analysis of population structure as previously described in Pritchard, Stephens,

126   & Donnelly (2000) was carried out to estimate the number of potential clusters present in the

127   microsatellite data, and to assign individuals to inferred clusters by STRUCTURE. Specifically,

128   five independent runs were carried for different values of K between 1 and 8, using no prior

129   information about individual location, and assuming admixture and correlated allele frequencies.

130   The Markov Chain Monte Carlo (MCMC) was run for a total of 1 million generations discarding

131   the first 100,000 as burn-in. The most likely K explaining the variation in the data was selected

132   estimating the maximal value of the log likelihood [Ln Pr(X/K)] of the posterior probability of

133   the data for a given K (Pritchard, Stephens, & Donnelly, 2000), and the ΔK statistic (Evanno,

134   Regnaut, & Goudet, 2005). The population structure results were graphically displayed by the

135   software DISTRUCT (Rosenberg, 2004). In addition, we visualized the genetic differentiation

136   among all samples with a factorial correspondence analysis (FCA) in GENETIX version 4.0.

137   Furthermore, we constructed a population graph network described by Dyer and Nason (2004)

138   using the popgraph package (Dyer, 2014) in R 2.15.3 (R Development Core Team 2013). The

139   method is based on the genetic covariance structure among populations analyzed simultaneously

140 (Dyer & Nason, 2004). Populations that exhibit significant genetic matrix correlation will be

141 connected in the network by edges (lines), and the length of the edges is inversely proportional to

142 the genetic covariance between the populations. Therefore, longer edges indicate lower genetic

143 covariance between populations. Populations that are not connected indicate the absence of

144 migration, and the presence of subgraphs (a smaller network within a large network) indicates

145 that a population or group of populations maintain a weak or null genetic connection (Dyer, 2007;

146 Dyer & Nason, 2004; Dyer, Nason, & Garrick, 2010).

147   Demographic history was performed in BOTTLENECK 1.2.02 (Piry, Luikart, & Cornuet,

148 1999) and assessed using Wilcoxon's sign rank test and mode-shift test as previously described

149 in Cornuet & Luikart (1996) and Luikart & Cornuet (1998), respectively. The software MSVAR

150 v.1.3 was used to characterize the recent demographic history of the whole *T. mongolica*

151 population based on the microsatellite data as described in Storz, Beaumont, & Alberts (2002).

152 Specifically, this method assumes that a current population (of size N_0) passed through a

153 demographic change (a bottleneck or an expansion) at time T in the past, which changed its size

154 from N_1 to N_0 following an exponential model. Five independent simulations were run to

155 estimate the distributions of these three parameters. For *T. mongolica*, the average generation

156 time is four years (Xu et al., 1998), and this period was adopted for the simulation. Each

157 MSVAR run consisted of 2×10^9 iterations of the MCMC algorithm discarding the first 10% of

158 the coalescent simulations as burn-in. The median (50%) of the posterior distributions were

159 calculated from five runs data. Finally, we plotted the marginal posterior distributions of the

160 three parameters by the LOCFIT package (Loader, 2007) implemented in R based on five runs.

RESULTS

Genetic diversity

In this study, a total of 339 individuals were genotyped at 12 loci. Micro-Checker did not indicate null alleles or genotyping errors such as large allele dropout or stuttering. There was no linkage disequilibrium at any locus in any population. The MNA for the eight populations varied between 13.17 and 17.67 with an overall value of 15.25. The overall observed heterozygosity (H_O) was 0.840 (0.810 − 0.873), while the overall expected heterozygosity (H_E) was 0.868 (0.832 − 0.882) (Table 1). Allelic richness (AR) ranged from 6.860 − 10.529, with an overall allelic richness across loci of 9.382 (Table 1). Inbreeding coefficient analysis generated negative values in SZS and QLS populations (Table 1). For the population as a whole, genetic diversity as characterized by microsatellite markers was higher than that reported for other shrub species (Table 2).

Population structure and genetic relationship

Based on STRUCTURE analysis, the Dealt K statistics output showed a clear maximum at $K = 2$ (Delta $K = 4.98$) (Fig. 2B), but no obvious maximum log likelihood of posterior probability was found (Ln P(K) = − 19881.82) (Fig. 2A). The ΔK value was remarkable at $K = 6$ (Delta $K = 3.69$) (Fig. 2B), with an obvious maximum log likelihood of posterior probability (Ln P(K) = − 19596.44) (Fig. 2A). These data suggest that six potential genetic clusters may exist among them. Notably, factors such as recent admixture, admixture with unsampled / unobservable "ghost" populations, and recent bottlenecks may lead to misinterpretation of STRUCTURE results (Gilbert et al., 2012; Lawson et al., 2012; Falush, Van Dorp & Lawson,

182   2016). According to this framework, $K = 6$ may be a pseudophase. The highly mixed color bars

183   in the DISTRUCT diagram (for $K = 2 - 6$, Fig. 2C) indicated strong admixture among the eight

184   populations. Furthermore, F_{ST} values among these populations ranged from 0.00034 to 0.04284,

185   indicating a weak genetic differentiation across them (Table 3). Besides, IBD tests detected no

186   significant correlation between geographical distances and genetic distance for the whole

187   sampling (r = 0.0608, p <= 0.3940). Furthermore, no separate groups were identified in the FCA

188   analysis (Fig. 3). Specifically, all populations were highly clumped and overlapped (Fig. 3). The

189   popgraph software produced a population network with no subgraphs (Fig. 4). Overall, the

190   population network exhibited high genetic connection among the cohorts, where each population

191   was connected to at least four other populations.

192   **Population demography**

193   In present study, there is no significant signal of recent bottlenecks in eight populations

194   under both TPM and SMM model whilst the mode-shift test also showed a normal L-shaped

195   distribution of allele frequencies. However, the MSVAR results showed the posterior distribution

196   of N_0 and N_1 did not overlap under exponential models, which indicates that the whole

197   population passed through a significant reduction in effective population size (Fig. 5A).

198   Statistically, the average medians of the posterior distributions were approximately 2.9652 for

199   log N_0, and approximately 4.7938 for log N_1 (Fig. 5A). Therefore, for the present *T. mongolica*

200   population, the current effective population size (N_0) was approximately 923, while the ancestral

201   effective population size (N_1) was approximately 62,214, showing an approximately 67-fold

202   populationdecrease. Furthermore, the medians of the posterior distribution log $T = 3.7368$ (Fig.

203   5B), indicate a recent population decline took place approximately 5,455 years ago.

204   **DISCUSSION**

205   **Genetic diversity**

206   Because of human overexploitation, *T. mongolica* has undergone a dramatic population

207   decline in past decades (Ge et al., 2011). However, our assessment of genetic variation based on

208   microsatellite data reveals high levels of genetic diversity in this species. In the population as a

209   whole, high microsatellite diversity was detected, with MNA, H_O and H_E values of 15.45, 0.84

210   and 0.868, respectively (Table 1). Based on inter-simple sequence repeats (ISSR) marker, this

211   species' average gene diversity was estimated to be 0.177 within populations (H_E), and the H_O

212   ranged from 0.213 to 0.305, with an average of 0.263 at the population level (Ge et al., 2003).

213   Compared with Ge et al., (2003) study, we determined there to be extremely high genetic

214   diversity based on nuclear microsatellites in this species. Therefore, SSR may represent a more

215   advantageous alternative to assess genetic diversity than ISSR in *T. mongolica*. However, SSR

216   may also be over-estimating genetic diversity as we detected higher genetic diversity in *T.*

217   *mongolica* compared with other shrub species such as *Zygophyllum xanthoxylon*, *Ziziphus celata*,

218   *Adiantum capillus-veneris*, *Grevillea macleayana*, *Arabidopsis lyrata*, *Calothamnus quadrifidus*,

219   *Myrtus communis* and *Schiedea adamantis* (Table 2). Generally, species with high genetic

220   diversity are members of large populations that were geographically widespread in recent history

221   (González et al., 1998). In this study, the high genetic diversity of *T. mongolica* may reveal the

222   large effective size of ancestral populations, as supported by the demographic analysis using

223   MSVAR. Furthermore, from a conservation perspective, it also implies that the recent sharp

224   population decline event did not have a significant effect on the genetic diversity of *T. mongolica*.

225   The conservation status of *T. mongolica* is however, clearly under severe threat and this study

226   indicates that urgent measures need to be put into place to ensure its ongoing survival.

227   **Population genetic structure**

228   Landscape features such as rivers and mountains can function as geographical barriers to

229   dispersal and gene flow, shaping population structure (Funk et al., 2005; Whiteley et al., 2004).

230   STRUCTURE analysis did not clearly identify genetic clusters corresponding to specific

231   populations (Fig. 2). Clustering results indicated unobstructed admixture and thus weak genetic

232   differentiation among *T. mongolica* populations. This result was corroborated by the pairwise

233   F_{ST} estimates and FCA analysis (Table 2, Fig. 3). Moreover, popgraph analysis showed that the

234   genetic structure is weak, and all of the samples were not assigned to any genetic group,

235   suggesting ongoing admixture processes between the extant populations (Fig. 4). For most

236   angiosperms, nuclear genes are inherited paternally via pollen, and maternally via seeds, while

237   cytoplasmic genes found in the chloroplast and mitochondria are maternally inherited (Petit,

238   Kremer & Wagner, 1993). Complex configurations of gene flow within and among populations

239   are expected through nuclear and chloroplast markers (Petit et al., 2005). In this study, the

240   patterns of genetic structure inferred from nuclear microsatellite markers suggests that the

241   Yellow River and Zhuozi Mountain do not act as significant barriers to pollination among

242   populations.

243   The Yellow River, the second longest river in China, is well-known for its frequent flooding

244   and heavy silt load (Sinclair, 1987). In the last 3,000 years, the river's levees have breached more

245   than 1,500 times and its course has changed approximately 26 times (Leung, 1996). As a result,

246   *T. mongolica* populations on the flood plain have been exposed to periodic habitat destruction

247   and fragmentation (Ge et al., 2011). It is established that species with narrow distributions and

248   small population sizes face a high risk of extinction, especially when gene flow between sub-

249   populations is restricted (Frankham, Ballou & Briscoe DA, 2002; Hanski & Gilpin 1997). In

250   seed plants, such gene flow occurs via the movement of pollen or seeds. Fortunately, *T.*

251   *mongolica* is primarily pollinated by insects (Zhen & Liu, 2008), negating the potential barrier

252   effect of the Yellow River and, to some degree, the Zhuozi Mountain (Fig. 1). Hence, neither

253   distinguishable genetic clusters nor population differentiation were detected in populations

254   separated by these barriers.

255   **Population demography**

256   In present study, neither heterozygosity excess nor mode-shift tests suggested a recent

257   population bottleneck for *T. mongolica*. However, MSVAR simulation indicated a severe recent

258   population decline in all populations (Fig. 5). Under the exponential model, the posterior

259   distribution of N_0 and N_1 (50% quantile) indicates a 67-fold population decline, starting

260   approximately 5,455 years ago, and is mirrored in similar declines for animals. For example, in

261   Northeastern Malaysia, human-induced deforestation and habitat fragmentation resulted in a

262   recent population collapse in orangutans, *Pongo pygmaeus*, approximately 210 years ago

263   (Goossens et al., 2006). Humans in southwestern China, over the course of thousands of years,

264   have caused the dramatic decline of the giant panda, *Ailuropoda melanoleuca* (Zhang et al., 2007)

265   and the tufted deer, *Elaphodus cephalophus* in the Yangtze River area (Sun et al., 2016). These

266 events suggest the possibility of an anthropogenically-induced decline for *T. mongolica*. In

267 addition, it is worth noting that the high-density human activities along the G6 road could also

268 significantly impact the *T. mongolica* populations in the next few decades.

269 **Implications for conservation**

270 Population genetics studies can help to identify management units (MUs) and evolutionarily

271 significant units (ESUs) for conservation (Moritz, 1994). In this study, all of the analytical

272 results indicate weak genetic differentiation among extant populations of *T. mongolica*. Our

273 work suggests that the eight *T. mongolica* populations sampled may be deemed a single MU for

274 conservation purposes. With rapidly increasing human disturbance, *T. mongolica* populations are

275 suffering from overexploitation, habitat loss and fragmentation, most noticeably along the G6

276 road. To better maintain the population size of *T. mongolica*, we propose that the Chinese

277 government should give greater priority to the conservation and restoration of its habitat, and to

278 plant more artificial populations in the core area of its current range along the G6 Road.

279 **CONCLUSION**

280 In this study, 339 individuals from eight populations were successfully genotyped at 12

281 nuclear loci, successfully. Based on microsatellite data, high levels of genetic diversity were

282 revealed in this endangered species. This study implies that the wild *T. mongolica* populations

283 still harbor a surprisingly rich gene pool. Furthermore, neither distinguishable genetic clusters

284 nor population differentiation were detected among extant *T. mongolica* populations. Finally, a

285 strong and recent population decline event was discovered, which was likely to have been

286 brought about by recent human activities, and emphasizes the need for urgent conservation

287 measures to ensure its ongoing survival.

288 **ACKNOWLEDGEMENTS**

289 The authors would like to thank Qiang Wang for his help in sample collecting. We would

290 also like to thank Dr. Jessica Pérez-Alquicira (National Autonomous University of Mexico,

291 Mexico), Dr. Tao Pan, Dr. Chaochao Hu and Ms. Guiyou Wu for their laboratory assistance and

292 insightful advice. We would also like to thank Mrs. Lise K. Sorensen (University of Utah, USA)

293 for her linguistic assistance.

294 **REFERENCES**

295 **Albaladejo RG, Sebastiani F, Gonzálezmartínez SC, Gonzálezvaro JP, Vendramin GG,**

296 **Aparicio A. 2010.** Isolation of microsatellite markers for the common mediterranean

297 shrub *Myrtus communis* (Myrtaceae). *American Journal of Botany* **97:** 23–25.

298 **Beaumont MA, Bruford MW. 1999.** *Microsatellites in conservation genetics*. Oxford: Oxford

299 University Press.

300 **Belkhir K, Borsa P, Chikhi L. 2001.** 1996-2001 GENETIX 4.02, logiciel sous Windows TM

301 pour la génétique des populations.

302 **Byrne M, Elliott CP, Yates C, Coates DJ. 2007.** Extensive pollen dispersal in a bird-pollinated

303 shrub, *Calothamnus quadrifidus*, in a fragmented landscape. *Molecular Ecology* **16:**

304 1303–1314.

305 **Clauss MJ, Mitchell-Olds T. 2006.** Population genetic structure of *Arabidopsis lyrata*, in

306 Europe. *Molecular Ecology* **15:** 2753–2766.

307 **Cornuet JM, Luikart G. 1996.** Description and power analysis of two tests for detecting recent

308 population bottlenecks from allele frequency data. *Genetics* **144:** 2001–2014.

309 **Culley TM, Weller SG, Sakai AK, Putnam KA. 2008.** Characterization of microsatellite loci

310 in the Hawaiian endemic shrub *Schiedea adamantis* (Caryophyllaceae) and

311 amplification in related species and genera. *Molecular Ecology Resources* **8:** 1081–1084.

312 **Dyer RJ, Nason JD, Garrick RC. 2010.** Landscape modelling of gene flow: Improved power

313 using conditional genetic distance derived from the topology of population networks.

314 *Molecular Ecology* **19:** 3746–3759.

315 **Dyer RJ, Nason JD. 2004.** Population Graphs: The graph theoretic shape of genetic structure.

316 *Molecular Ecology* **13:** 1713–1727.

317 **Dyer RJ. 2007.** The evolution of genetic topologies. Theoretical Population Biology **71:** 71–79.

318 **Dyer RJ. 2014.** R package gstudio: analyses and functions related to the spatial analysis of

319 genetic marker data. Retrieved from https://github. com/dyerlab/gstudio.git

320 **Dong XJ, Zhang XS. 2001.** Some observations of the adaptations of sandy shrubs to the arid

321 environment in the MU Us Sandland: leaf water relations and anatomic

322 features. *Journal of Arid Environments* **48:** 41–48.

323 **Doyle JJ, Doyle JL. 1987.** A rapid isolation procedure for small quantities of fresh leaf tissue.

324 *Phytochem Bull* **19:**11–15.

325 **Earl DA, VonHoldt BM. 2012.** STRUCTURE HARVESTER: a website and program for

326 visualizing STRUCTURE output and implementing the Evanno mehod. *Conservation*

327 *Genetics Resources* **4**: 359–361.

328 **Eckstein RL, O'Neill RA, Danihelka J, Otte A, Köhler W. 2006.** Genetic structure among and

329 within peripheral and central populations of three endangered floodplain

330 violets. *Molecular Ecology* **15:** 2367–2379.

331 **England PR, Usher AV, Whelan RJ, Ayre DJ. 2002.** Microsatellite diversity and genetic

332 structure of fragmented populations of the rare, fire-dependent shrub *Grevillea*

333 *macleayana*. *Molecular Ecology* **11:** 967–977.

334 **Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using

335 the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

336 **Excoffier L, Laval G, Schneider S. 2005.** Arlequin (version 3.0): an integrated software

337   package for population genetics data analysis. *Evolutionary Bioinformatics* **1:** 47–50.

338 **Falush D, van Dorp L, Lawson D. 2016.** A tutorial on how (not) to overinterpret

339   STRUCTURE/ADMIXTURE bar plots. *bioRxiv* 66431.

340 **Frankham R, Ballou JD, Briscoe DA. 2002.** *Introduction to Conservation Genetics.* Cambridge:

341   Cambridge University Press.

342 **Fu LG. 1992.** *The Red Book of Chinese Plants–Rare and Endangered Plants.* Beijing: Science

343   Press.

344 **Funk CW, Blouin MS, Corn PS Maxell BA, Pilliod DS, Amish S, Allendorf FW. 2001.**

345   Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected

346   by the landscape. *Molecular Ecology* **14:** 483–496.

347 **Gaines MS, Diffendorfer JE, Tamarin RH, Whittam TS. 1997.** The effects of habitat

348   fragmentation on the genetic structure of small mammal populations. *Journal of*

349   *Heredity* **88:** 294–304.

350 **Ge XJ, Hwang CC, Liu ZH, Huang CC, Huang WH, Hung, KH, Wang WK, Chiang TY.**

351   **2011.** Conservation genetics and phylogeography of endangered and endemic shrub

352   *Tetraena mongolica* (Zygophyllaceae) in Inner Mongolia, China. *BMC Genetics* **12:** 1.

353 **Ge XJ, Yu Y, Zhao NX, Chen HS, Qi WQ. 2003.** Genetic variation in the endangered Inner

354   Mongolia endemic shrub *Tetraena mongolica* Maxim (Zygophyllaceae). *Biological*

355   *Conservation* **111:** 427–434.

356 **Gilbert KJ, Andrew RL, Dan GB, Franklin MT, Kane NC, Moore JS, Moyers BT, Renaut**

357   **S, Rennison DJ, Veen T. 2012.** Recommendations for utilizing and reporting

358   population genetic analyses: the reproducibility of genetic clustering using the program

359   structure. *Molecular Ecology* **21:** 4925–4930.

360   **Gitzendanner MA, Weekley CW, Germain-Aubrey CC, Soltis DE, Soltis PS. 2012.**

361   Microsatellite evidence for high clonality and limited genetic diversity in *Ziziphus*

362   *celata*, (Rhamnaceae), an endangered, self-incompatible shrub endemic to the Lake

363   Wales Ridge, Florida, USA. *Conservation Genetics* **13:** 223–234.

364   **González S, Maldonado JE, Leonard JA, Vil C, Barbanti Duarte JM, Merino M, Brum-**

365   **Zorrilla N, Wayne RK. 1998.** Conservation genetics of the endangered Pampas deer

366   (*Ozotoceros bezoarticus*). *Molecular Ecology* **7:** 47–56.

367   **Goossens B, Chikhi L, Ancrenaz M, Lackman-Ancrenaz I, Andau P, Bruford MW. 2006.**

368   Genetic signature of anthropogenic population collapse in orang-utans. *PLoS Biology* **4:**

369   e25.

370   **Goudet J. 2002.** *FASTA, a program to estimate and test gene diversities and fixation indices.*

371   Lausanne: Institute of Ecology.

372   **Hanski IA, Gilpin ME. 1997.** *Metapopulation biology: ecology, genetics, and evolution.*

373   Orlando: Academic Press.

374   **Holland MM, Parson W. 2011.** GeneMarker(R) HID: A reliable software tool for the analysis

375   of forensic STR data. *Journal of Forensic Sciences* **56:** 29–35.

376   **Jensen JL, Bohonak AJ, Kelley ST. 2005.** Isolation by distance, web service. *BMC Genetics* **6:**

377   13.

378   **Lawson DJ, Hellenthal G, Myers S, Falush D. 2012.** Inference of population structure using

379   dense haplotype data. *Plos Genetics* **8:** e1002453

380   **Leung G. 1996.** Reclamation and Sediment Control in the Middle Yellow River Valley. *Water*

381   *International.* **21:**12–19.

382   **Loader C. 2007.** Locfit: Local regression, likelihood and density estimation. R package version

383   1–5.

384   **Luikart G, Cornuet JM. 1998.** Empirical evaluation of a test for identifying recently

385   bottlenecked populations from allele frequency data. *Conservation Biology* **12:** 228–237.

386

387   **Manel S, Schwartz MK, Luikart G, Taberlet P. 2003.** Landscape genetics: combining

388   landscape ecology and population genetics. *Trends in Ecology & Evolution* **18:** 189–197.

389   **Moritz C. 1994.** Defining 'evolutionarily significant units' for conservation. *Trends in Ecology*

390   *& Evolution* **9:** 373–375.

391   **Nupp TE, Swhart RK. 1998.** Effects of forest fragmentation on population attributes of white

392   footed mice and eastern chipmunks. *Journal of Mammalogy* **79:** 1234–1243.

393   **O'Brien SJ. 1994.** A role for molecular genetics in biological conservation. *Proceedings of the*

394   *National Academy of Sciences of the USA* **91:** 5748–5755.

395   **Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. 2005.** Comparative

396   organization of chloroplast, mitochondrial and nuclear diversity in plant populations.

397   *Molecular Ecology* **14:** 689–701.

398   **Petit RJ, Kremer A, Wagner DB. 1993.** Finite island model for organelle and nuclear genes in

399 plants. *Heredity* **71:** 630–641.

400 **Piry S, Luikart G, Cornuet JM. 1999.** BOTTLENECK: a computer program for detecting

401 recent reductions in the effective size using allele frequency data. *Journal of Heredity* **90:**

402 502–503.

403 **Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using

404 multilocus genotype data. *Genetics* **155:** 945–959.

405 **Pryor KV, Young JE, Rumsey FJ, Edwards KJ, Bruford MW, Rogers HJ. 2001.** Diversity,

406 genetic structure and evidence of outcrossing in british populations of the rock fern

407 *Adiantum capillus-veneris*, using microsatellites. *Molecular Ecology* **10:** 1881–1894.

408 **R Development Core Team. 2013.** R: A language and environment for statistical computing.

409 Vienna: R Foundation for Statistical Computing.

410 **Rockwood LL. 2006.** *Introduction to population ecology.* Oxford: Blackwell Publishing Ltd.

411 **Rosenberg NA. 2004.** DISTRUCT: a program for the graphical display of population structure.

412 *Molecular Ecology Notes* **4:** 137–138.

413 **Rousset F. 2008.** GENEPOP'007: a complete re-implementation of the genepop software for

414 Windows and Linux. *Molecular Ecology Resources* **8:** 103–106.

415 **Rousset F, Raymond, M. 1997.** Statistical analyses of population genetic data, new tools, old

416 concepts. *Trends in Ecology & Evolution* **12:** 313–317.

417 **Sinclair K. 1987.** *The Yellow River: A 5000 Year Journey Through China Chatswood.* Sydney:

418 Child & Associates Publishing.

419 **Song G, Yu LJ, Gao B, Zhang RY, Qu YH, Lambert DM, Li SS, Zhou TL, Lei FM. 2013.**

420   Gene flow maintains genetic diversity and colonization potential in recently range-

421   expanded populations of an oriental bird, the Light-vented Bulbul (*Pycnonotus sinensis*,

422   Aves: Pycnonotidae). *Diversity and Distributions* **19:** 1248–1262.

423   **Storz JF, Beaumont MA, Alberts SC. 2002.** Genetic evidence for long-term population decline

424   in a savannah-dwelling primate: inferences from a hierarchical Bayesian model.

425   *Molecular Biology and Evolution* **19:** 1981-1990.

426   **Sun ZL, Pan T, Wang H, Pang MJ, Zhang BW. 2016**. Yangtze River, an insignificant genetic

427   boundary in tufted deer (*Elaphodus cephalophus*): the evidence from a first population

428   genetics study. *PeerJ* **4:** e2654.

429   **Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004.** MICRO-CHECKER:

430   software for identifying and correcting genotyping errors in microsatellite data.

431   *Molecular Ecology Notes* **4:** 535–538.

432   **Wang YC, Ma H, Zheng R. 2000.** Studies on the reproductive characteristics of *Tetraena*

433   *mongolica* Maxim. *Acta Botanica Boreal Occident Sinica* **20:** 661–665

434   **Whiteley AR, Spruell P, Allendorf FW. 2004.** Ecological and life history characteristics

435   predict population genetic divergence of two salmonids in the same landscape.

436   *Molecular Ecology* **13:** 3675–3688.

437   **Xu LR, Huang CJ, Huang BX, Liu YX, Li BT, Zhang YT. 1998.** *Flora Republicae Popularis*

438   *Sinicae* (vols.43). Beijing: Science Press.

439   **Zhang BW, Li M, Zhang ZJ, Goossens B, Zhu LF, Zhang SN, Hu JC, Bruford MW, Wei**

440   **FW. 2007.** Genetic viability and population history of the giant panda, putting an end to

441   the "evolutionary dead end"? *Molecular Biology and Evolution* 24: 1801–1810.

442   **Zhang YF, Yang C, Li B, Chen JK. 2003.** Interpopulation difference in growth and

443   reproduction of endemic species *Tetraena mongolica* in Ordos Plateau. *Acta Ecologica*

444   *Sinica* **23:** 436–443.

445   **Zhang YJ, Yang C. 2000.** Comparative analysis of genetic diversity in the endangered shrub

446   *Tetraena mongolica* and its related congener *Zygophyllum xanthoxylon*. *Acta*

447   *Phytoecologica Sinica* **24:** 425–429.

448   **Zhen JH, Liu GH. 2008.** Research advance in rare and endemic plant *Tetraena mongolica*

449   Maxim. *Chinese Journal of Applied Ecology* **19:**433–440.

450   **Zhi YB, Ding L, Wang YP, Hu LF, Hua YP, Zhao K, He JQ, Yang C, Zhang BW. 2014.**

451   Isolation and characterization of thirteen novel dinucleotide microsatellite loci from

452   *Tetraena mongolica* Maxim. *Conservation Genetics Resources* **6:** 297–299.

Figure 1

Map showing the population location of *T. mongolica* sampled in this study.

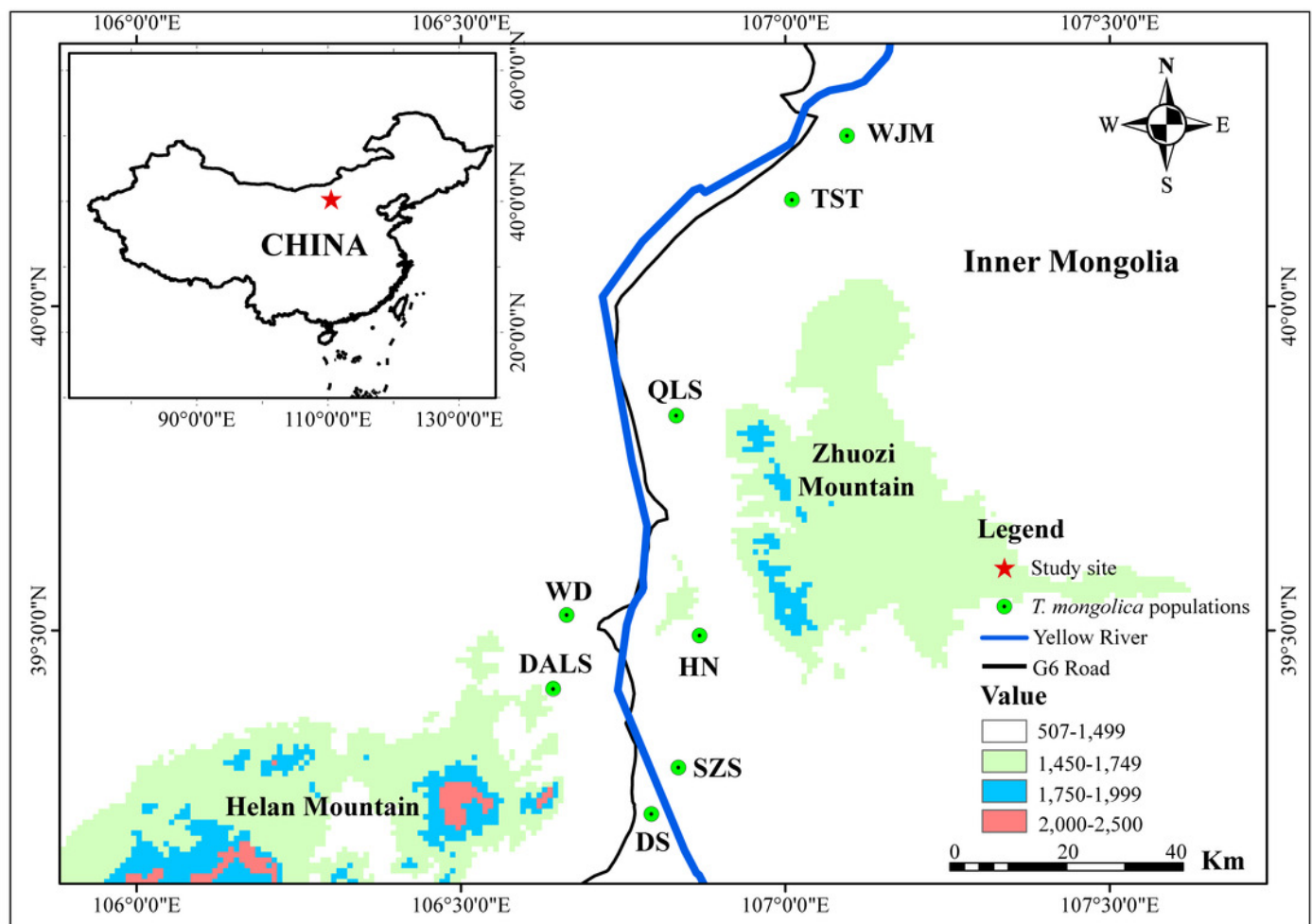



Figure 2

Bayesian STRUCTURE clustering results of microsatellite variation among *T. mongolica* populations.

(A) The linear relationship between LnP(D) and *K*, (B) Dealt *K* values as a function of *K* based on 5 runs and (C) STRUCTURE output from *K* = 2 to 6.

(A) L(*K*) (mean +- SD)

(B) Dealt *K* = mean (L''(*K*)) / sd (L(*K*))

(C)

K=2

K=3

K=4

K=5

K=6

SZS   DS   HN   DALS   WD   QLS   WJM   TST

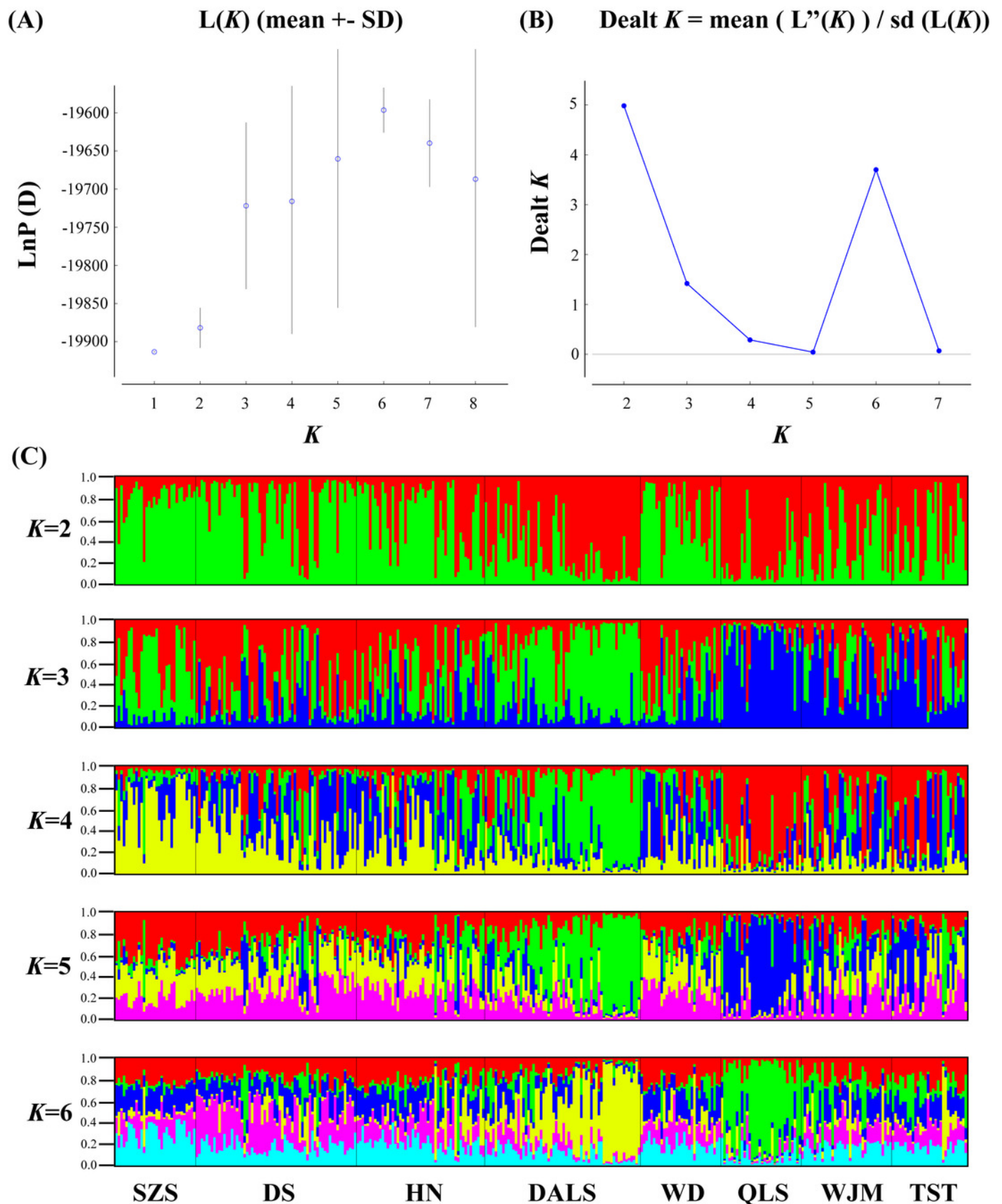



Figure 3

Factorial correspondence analysis performed for *T. mongolica* based on nuclear microsatellite loci.

Symbols and colors represent individuals from different populations.

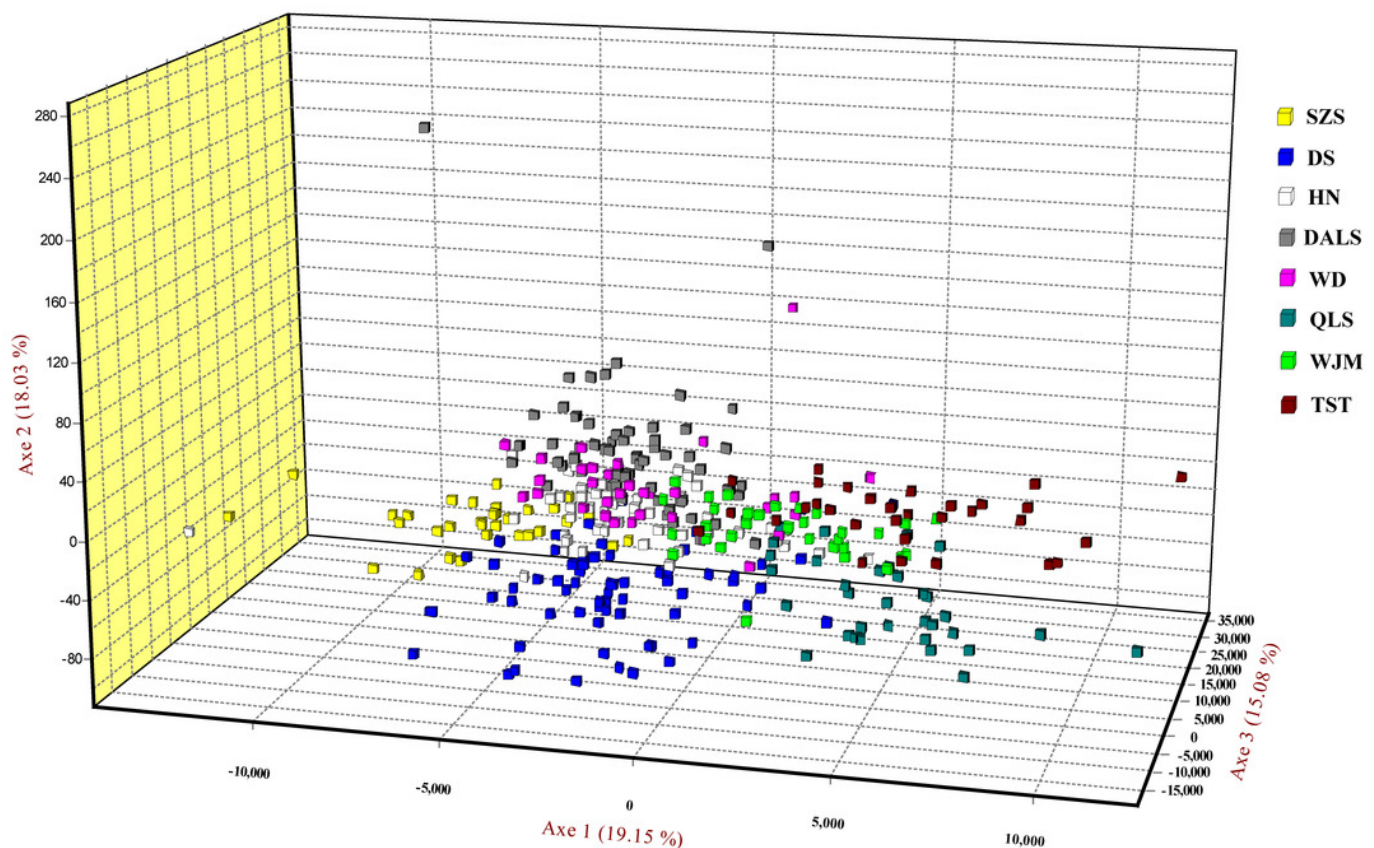



Figure 4

Population graph for 8 populations of *Tetraena mongolica* Maxim based on nuclear microsatellites data.

The size of the nodes (spheres) represents the genetic variation within populations and edges (lines) connect directly two populations showing significant genetic covariance.

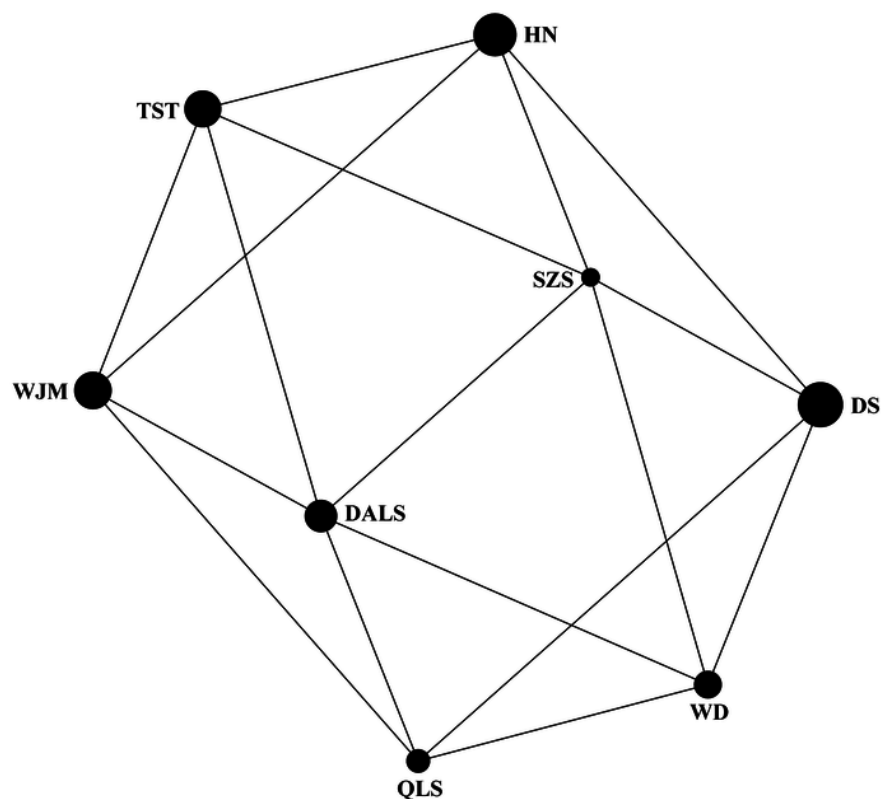
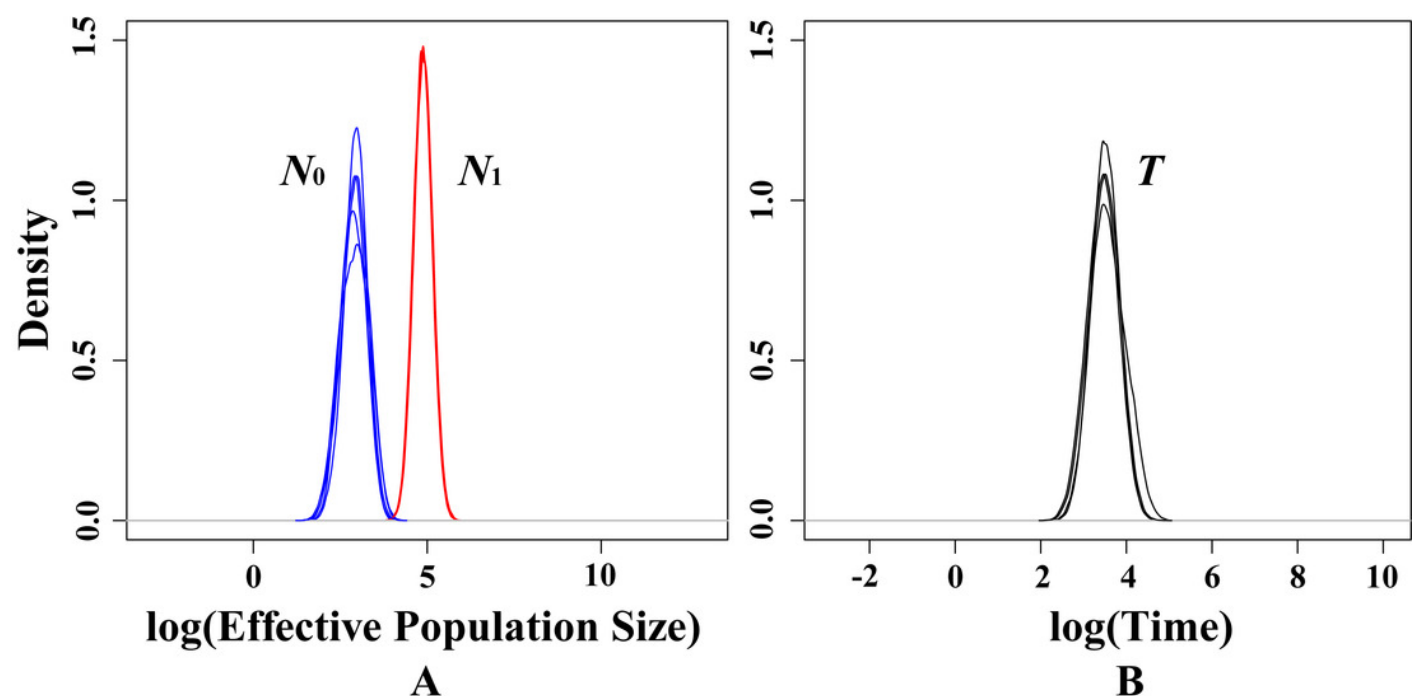



Figure 5

Estimated posterior distributions of N_0, N_1 and T using MSVAR.

N_0, current effective population sizes (blue curve); N_1, ancestral effective population sizes (red curve); T, time since population change (black curve). All densities are represented in a log10 scale.

Genetic variability observed within populations using microsatellite loci.

Note: N, number of individuals; MNA, mean number of allele per locus; AR, allelic richness; H_o and H_E, observed and expected heterozygosity; F_{IS}, inbreeding coefficient.

1 **Table 1. Genetic variability observed within populations using smicrosatellite loci.**

Population	N	MNA	AR	H_O	H_E	F_{IS} (IC 95%)
SZS	32	13.17	6.860	0.850	0.832	-0.00163 (-0.03378 - 0.00409)
DS	64	17.67	8.284	0.836	0.881	0.06051 (0.02421 - 0.07440)
NH	51	17.67	10.499	0.812	0.880	0.08889 (0.04226 - 0.09998)
DALS	62	16.75	8.974	0.810	0.867	0.07559 (0.04021 - 0.09279)
WD	32	14.67	8.829	0.851	0.867	0.03657 (-0.02664 - 0.07773)
QLS	32	13.92	8.532	0.873	0.854	-0.00596 (-0.05695 - 0.00596)
WJM	36	15.67	10.529	0.852	0.882	0.04891 (-0.01434 - 0.05576)
TST	30	14.08	9.397	0.840	0.877	0.06018 (0.02117 - 0.06018)
Total	339	15.45	9.382	0.840	0.868	0.06760 (0.05360 - 0.06888)

2 Note: N, number of individuals; MNA, mean number of allele per locus; AR, allelic richness; H_o and H_E, observed and expected heterozygosity; F_{IS},

3 inbreeding coefficient.

Genetic diversity of *Tetraena mongolica* and other shrub based on microsatellite loci.

1 **Table 2. Genetic diversity of *Tetraena mongolica* and other shrub based on nuclear microsatellite loci.**

Species	N	MNA	H_O	H_E	Reference
Tetraena mongolica	339	15.45	0.84	0.868	In this study
T. mongolica	338	1.6	0.199	0.345	Zhang and Yang 2000
Zygophyllum xanthoxylon	61	2.2	0.43	0.392	Zhang and Yang 2000
Ziziphus celata	595	2.23	0.69	0.39	Gitzendanner et al., 2001
Adiantum capillus-veneris	151	-	0.13-0.37	0.2-0.63	Pryor et al., 2001
Grevillea macleayana	321	-	0.248-0.523	0.420-0.523	England et al., 2002
Arabidopsis lyrata	344	9.3	0.48	0.52	Clauss and Mitchell-Olds 2006
Calothamnus quadrifidus	114	19.67	0.584	0.867	Byrne et al., 2007
Myrtus communis	48	-	0.258-0.802	0.125-0.875	Albaladejo et al., 2010
Schiedea adamantis	49	-	0.125-0.755	0.041-0.787	Culley et al., 2008

Pairwise F_{ST} estimates based on nuclear microsatellite loci.

Note: The asterisks (*) mean $P<0.05$.

1   **Table 3. Pairwise F_{ST} estimates based on nuclear microsatellite loci.**

Populations	1	2	3	4	5	6	7	8
1. SZS								
2. DS	0.01599*							
3. HN	0.00607*	0.00620*						
4. DALS	0.01648*	0.01623*	0.0076*					
5. WD	0.01795*	0.00834*	0.00793*	0.01522*				
6. QLS	0.04284*	0.02743*	0.01740*	0.02972*	0.03673*			
7. WJIM	0.01580*	0.00944*	0.00034	0.00769*	0.01095*	0.01839*		
8. TST	0.02699*	0.01837*	0.01373*	0.01829*	0.02349*	0.02138*	0.00912*	

2   Note: The asterisks (*) mean $P<0.05$.