RESEARCH ARTICLE

Molecular phylogeography of East Asian *Boea clarkeana* (Gesneriaceae) in relation to habitat restriction

Ying Wang[1,2], Kun Liu[1,2], De Bi[1], Shoubiao Zhou[1,3]*, Jianwen Shao[1,2]*

1 College of Life Sciences, Anhui Normal University, Wuhu, Anhui, China, **2** Anhui Provincial Key Laboratory of the Conservation and Exploitation of Biological Resources, Wuhu, Anhui, China, **3** College of Environmental Science and Engineering, Anhui Normal University, Wuhu, Anhui, China

* zhoushoubiao@vip.163.com (SZ); 545491044@qq.com (JS)

Abstract

Subfamily Cyrtandroideae (Gesneriaceae) comprises a broadly distributed group of rocky-slope herbs, with China being the center of its distributional range. The normal growth of many species within the family is particularly dependent on special habitats. Due to the paucity of molecular studies, very little is known regarding East Asian herb phylogeographic pattern. Here, we investigate the molecular phylogeography of *Boea clarkeana* Hemsl., a unique resurrection herb endemic to China, focusing on geographically restrictive effects of habitat distribution on evolutionary history. Variation in three chloroplast DNA (cpDNA) intergenic spacers (*psb*A-*trn*H, *rps*12-*rpl*20, and *trn*L-*trn*F), the ribosomal internal transcribed spacer (ITS) and simple sequence repeats in expressed sequence tags (EST-SSRs) was investigated across 18 populations to assess genetic diversity, genetic structure and historical dynamics. Genetic diversity was low within populations (cpDNA, $h_S = 0.03$, $\pi_S \times 10^3 = 0.17$; ITS, $h_S = 0.16$, $\pi_S \times 10^3 = 0.43$) but high for species (cpDNA, $h_T = 0.82$, $\pi_T \times 10^3 = 3.12$; ITS, $h_T = 0.88$, $\pi_T \times 10^3 = 6.39$); 76 alleles were detected in this highly inbred species ($F_{IS} = 0.22$), with a significantly low average of 1.34 alleles per locus. No cpDNA or ITS haplotypes were shared between regions. Based on cpDNA results, the Mt. Huangshan-Tianmu and Mt. Qinling-Daba haplotypes are ancestral; these two regions represent potential refugia. Although no evidence of significant retreat-migration phenomena during glacial cycles was detected, interglacial range expansion from northern Mt. Qinling-Daba was identified (121,457 yr BP). Rapid agricultural growth caused bottlenecks in many populations, especially on Mt. Huang-Tianmu. Habitat restriction and fragmentation, weak seed and pollen dispersal abilities, and long-term isolation caused by human-induced or environmental changes are considered the main causes of extinction of several populations and low genetic diversity within populations and regions. These analyses clarify the effects of habitat restriction on *B. clarkeana*, representing an evolutionary reference for similar gesneriads, and enrich our understanding of the molecular phylogeography of East Asian rocky-slope herbs.

Introduction

A stable climate, mature plant communities in vast landscapes, high spatial heterogeneity and a long evolutionary history have laid the foundation for floristic richness in China. Accordingly, this area is an important region of diversity in the Northern Hemisphere (with a very high level of endemism [1]) and an important center for species conservation, speciation and evolution [2]. However, molecular phylogeographic studies of this area have largely concentrated on woody (mostly relic) plants [3–12]. Compared with long-lived woody plants, herbs undergo far more life cycles within a given time period and are therefore expected to respond more quickly to environmental changes; thus, herbaceous plants may provide better opportunities to study the drivers of diversification and speciation [13, 14]. However, to our knowledge, only a few herbs in this area, such as *Dysosma versipellis*, *Primula ovalifolia* and Chinese weedy rice (*Oryza stavia* L. f. sp. *spontanea*), have been studied to date [13, 15–17].

Gesneriaceae is a family of aesthetically appealing herbs with high ornamental value, among which the center of the distributional range of subfamily Cyrtandroideae occurs in China [18]. As a group of rocky-slope herbs, many species within the family have special requirements regarding soil types and habitats and are generally found in narrow areas [19]. Due to the many rare and endangered species in this family as well as a number of resurrection plants, which are extremely rare among other angiosperm groups [20], Gesneriaceae has attracted widespread attention from researchers [19, 21]. However, previous studies on Gesneriaceae have mainly focused on aspects such as population distribution, ecological protection, cytology, and ornamental value [19] or on molecular mechanisms and recovery processes (e.g., in *Boea crassifolia* and *Boea hygrometrica* [22–25]). In contrast, there have been few studies involving DNA sequences [26–28], and no research is available regarding the molecular phylogeography of these plants, even though such information may aid in understanding genetic structures and geographical distribution patterns in relation to the intrazonal distribution of Gesneriaceae.

Boea clarkeana Hemsl., a small perennial plant endemic to China, has remarkable light-blue flowers that confer high horticultural value [18]. *Boea clarkeana* is a typical resurrection plant of Gesneriaceae that exhibits excellent desiccation tolerance (DT) due to the ability to recover even after most of its cellular water has been lost for a period of time [20]. As a valuable gene pool for DT, *B. clarkeana* presents great research value, especially regarding the molecular mechanisms of DT, which may be helpful for crop cultivation [29]. Compared with other species of Gesneriaceae, *B. clarkeana* often shows restricted growth on rock outcrops (mostly poor limestone) in the shade of shrubs and trees [30] and has a much wider distribution area [18]. It is currently mainly distributed in China in a long, pear-shaped area along the middle-lower reaches of the Yangtze River [18, 31]. The genetic connectivity between populations of this plant mainly depends on entomophilous pollination and windborne seed dispersal. In recent years, suitable habitats have become more fragmented, and the numbers and sizes of populations have decreased significantly due to interference from human activities. Although *B. clarkeana* has become a subject of scientific interest in the last few years [30, 32, 33], its genetic diversity and structure, which are basic characteristics related to protection and utilization of the species, remain unknown.

Here, the population biodiversity and demographic history of *B. clarkeana* are examined based on variations in simple sequence repeats, or microsatellites, in expressed sequence tags (EST-SSRs) and specific regions of cpDNA (intergenic spacers, IGS) and nrDNA (entire internal transcribed spacer, ITS) in samples from 18 populations. The glacial refugia (during the last glacial maximum, LGM) and colonization routes in the postglacial (expansion) period for this plant were identified according to the results of these analyses. Overall, the results of this

study enable a better understanding of the influence of habitat restriction on *B. clarkeana*, which represents a good evolutionary reference for similar plant species of Gesneriaceae. Furthermore, this study enriches the published research regarding the molecular phylogeography of rocky-slope herbs of East Asia.

Materials and methods

Sampling, experimental design and DNA extraction

Samples were collected by and under the permit of the JiangXi Guanshan National Nature Reserve and the Protection of Wildlife Management Office of Anhui Province. A total of 394 individuals were collected from 18 natural populations covering the vast majority of the geographical range of *B. clarkeana* (see Table 1) [18, 31]. These samples were divided into four regional groups according to geographical distribution: Mt. Huang, Mt. Tianmu and Mt. Guan formed the eastern region; the central region included Mt. Shennongjia, Mt. Wu and

Table 1. Details of the samples, genetic diversity and haplotype composition of the *B. clarkeana* populations surveyed, based on cpDNA (*psb*A-*trn*H, *rps*12-*rpl*20, *trn*L-*trn*F), ITS and EST-SSR sequences.

Code	Location[a]	Longitude (˚E)	Latitude (˚N)	Size	EST-SSR[b]					cpDNA			ITS		
					N_A	PIC	H_O	H_E	F_{IS}	Haplotype	h (±SD)	$\pi \times 10^3$ (±SD×10³)	Haplotype	h (±SD)	$\pi \times 10^3$ (±SD×10³)
ZSC	T	120˚79'	29˚40'	21	1.533	0.1	0.093	0.118	0.373	C1,C2	0.15 (±0.13)	0.08 (±0.07)	R1,R2	0.15 (±0.13)	0.24 (±0.19)
ZTC	T	119˚44'	30˚32'	18	1.267	0.076	0.022	0.096	0.779	C1	0	0	R3	0	0
ZLQ	T	118˚90'	30˚12'	24	1.267	0.067	0.056	0.082	0.338	C2	0	0	R2	0	0
ANS	H	118˚88'	30˚68'	22	1.667	0.085	0.025	0.093	0.744	C2	0	0	R3	0	0
ANY	H	118˚13'	30˚89'	24	1.333	0.092	0.125	0.117	-0.051	C2	0	0	R1,R2,R4	0.39 (±0.15)	0.78 (±0.33)
ACT	H	117˚63'	30˚34'	24	1.533	0.095	0.111	0.111	0.017	C2	0	0	R1	0	0
AQL	H	117˚49'	29˚88'	24	1.4	0.088	0.106	0.105	0.012	C2	0	0	R1	0	0
JGD	G	114˚56'	28˚54'	24	1.133	0.008	0.008	0.008	-0.015	C3	0	0	R5	0	0
HSP	S	110˚85'	31˚75'	20	1.267	0.079	0.074	0.09	-0.043	C4	0	0	R6,R7,R8	0.67 (±0.08)	1.51 (±0.20)
HXS	S	110˚64'	31˚34'	14	1.007	0.005	0.005	0.005	0	C4	0	0	R8	0	0
HZS	W	110˚40'	30˚84'	24	1.4	0.068	0.036	0.082	0.607	C4	0	0	R8	0	0
HZL	Z	110˚45'	29˚09'	24	1.333	0.079	0.089	0.096	0.101	C5	0	0	R9,R10	0.57 (±0.12)	3.54 (±0.74)
HYJ	Z	110˚44'	28˚78'	24	2.333	0.27	0.278	0.309	0.121	C5	0	0	R11	0	0
CNJ	N	107˚20'	29˚10'	24	1.333	0.037	0.014	0.043	0.688	C6	0	0	R12	0	0
SNX	D	106˚93'	32˚44'	24	1.13	0.024	0.023	0.028	0.193	C7,C8	0.41 (±0.13)	0.22 (±0.13)	R13,R14	0.53 (±0.08)	1.64 (±0.24)
SWM	D	106˚54'	32˚57'	20	1.067	0.003	0.003	0.003	0	C7	0	0	R13	0	0
SLB	Q	106˚08'	33˚58'	24	1.133	0.023	0.029	0.027	-0.048	C7	0	0	R13	0	0
SLG	Q	105˚94'	33˚30'	15	1	0	0	0	NA	C7	0	0	R13	0	0
AVERAGE					1.34	0.07	0.06	0.08	0.22		0.03	0.17		0.16	0.43
TOTAL				394						8	0.82 (±0.01)	3.12 (±0.13)	14	0.88 (±0.01)	6.39 (±0.30)

Notes

[a]: H, Huang Mts.; T, Tianmu Mts.; G, Guan Mts.; W, Wu Mts.; S, Shennongjia Mts.; Z, Zhangjiajie Mts.; N, Nan Mts.; Q, Qinling Mts.; D, Daba Mts.

[b], N_A, number of alleles per locus across all populations; PIC, polymorphic information content; H_O, observed heterozygosity (mean value); H_E, expected heterozygosity (mean value); F_{IS}, inbreeding coefficient; h haplotype diversity; π, nucleotide diversity; SD, standard deviation; NA, not available.

Mt. Zhangjiajie; Mt. Qinling and Mt. Daba comprised the northwest region; and the southwest region included only one population (designated CNJ), located on Mt. Nan, the western-stretching branch of Mt. Wuling (see Fig 1A).

Leaf material was collected and dried rapidly in silicagel. All individuals (n = 394) were surveyed for population structure and genetic diversity with respect to EST-SSR variation, and cpDNA (intergenic spacer, IGS) and nrDNA (entire internal transcribed spacer, ITS) regions were sequenced in a subset of samples (n = 222). Total genomic DNA was isolated from dried leaves using QIAGEN DNeasy® Plant Mini Kit (QIAGEN, Germany) following the manufacturer's manual.

Amplification, sequencing and microsatellite genotyping

Polymerase chain reaction (PCR) amplification was performed using a 5828R PCR instrument (MyCycler Thermal Cycler, BIO-RAD, USA) in a 15-μL volume containing 50 ng DNA, 1×PCR buffer, 2.5 mM deoxyribonucleotide triphosphates (dNTPs) and MgCl$_2$, 0.5 U *Taq* polymerase (TaKaRa, Dalian, China), and 0.04 μM forward and reverse primers. The PCR conditions were as follows: initial denaturation at 94°C for 5 min, followed by 35 cycles of 30 s at 94°C, 40 s at 54°C, and 45 s at 72°C, with an additional extension at 72°C for 10 min. The PCR products were detected using an ABI 3730XL DNA Analyzer (Applied Biosystems, Foster City, USA), and alleles were scored manually using GeneMaker (version 2.2.0).

For cpDNA, six IGS regions (*psb*A-*trn*H, *rps*12-*rpl*20, *trn*L-*trn*F, *trn*S-*trn*G, *trn*T-*trn*F and *psb*B-*psb*F) were sequenced and detected in a preliminary screen [34, 35], after which three regions (*psb*A-*trn*H, *rps*12-*rpl*20 and *trn*L-*trn*F) showing stable amplification and high levels of variation were chosen for further analyses. Sequence alignments were generated using MEGA 6.0 [36]. Seventeen EST-SSR markers (BC1-BC17) that were specifically developed for *B. clarkeana* were used to detect genetic variation [32]. In the survey of EST-SSR variation, fluorescent M13 primers were used to accurately screen variation among individuals. All haplotype cpDNA and ITS sequences have been deposited in the GenBank database (*psb*A-*trn*H, BankIt1898693, KU867796-KU867801; *rps*12-*rpl*20, BankIt1898703, KU867802-KU867805; *trn*L-*trn*F, BankIt1898706, KU867806-KU867811; ITS, BankIt1898665, KU867781-KU867795).

Genetic diversity and structural analyses

The genetic diversity of the cpDNA and ITS sequences was estimated based on the number of haplotypes, haplotype diversity and nucleotide diversity for the species (h_T, π_T) and populations (h_S, π_S), which were determined using DnaSP 5.1 [37]. Phylogeographic structure was evaluated using the differentiation indices N_{ST} (ordered haplotypes) and G_{ST} (unordered haplotypes), which were calculated using Permut 1.0 (1000 permutations) [38]. To detect the correlation between genetic and geographical distance and estimate genetic variance components, the Mantel test and analysis of molecular variance (AMOVA) were implemented in Arlequin 3.0 [39, 40]. Neighbor-joining (NJ) phylogenetic trees of 8 chlorotypes and 14 ribotypes of *B. clarkeana* were built using Mega 6.0 [36].

For EST-SSRs, Hardy-Weinberg equilibrium (HWE) was detected using the online tool GENEPOP (http://www.genepop.curtin.edu.au/), and Bonferroni's correction was applied. We then utilized LOSITAN [41] for neutral detection of markers. Finally, 15 pairs of primers showing the greatest numbers of polymorphic loci, good amplification and agreement with the neutral theory (BC2 and BC14 did not agree with the neutral theory) were selected for the genotyping of all samples. The genetic variation of EST-SSRs was estimated based on the number of alleles (N_A), observed heterozygosity (H_O), expected heterozygosity (H_E) and polymorphism information content (PIC) at each locus, which were calculated using GenAlEx v.6.501 [42] and PowerMarker (version 3.25) [43]. The within-population inbreeding coefficient (F_{IS}) and

among-population coefficient (F_{ST}) were determined with FSTAT [44], and AMOVA was implemented in Arlequin 3.0 [39]. To reveal population structure based on the genetic distance among all sampled individuals, principal component analysis (PCoA) was performed
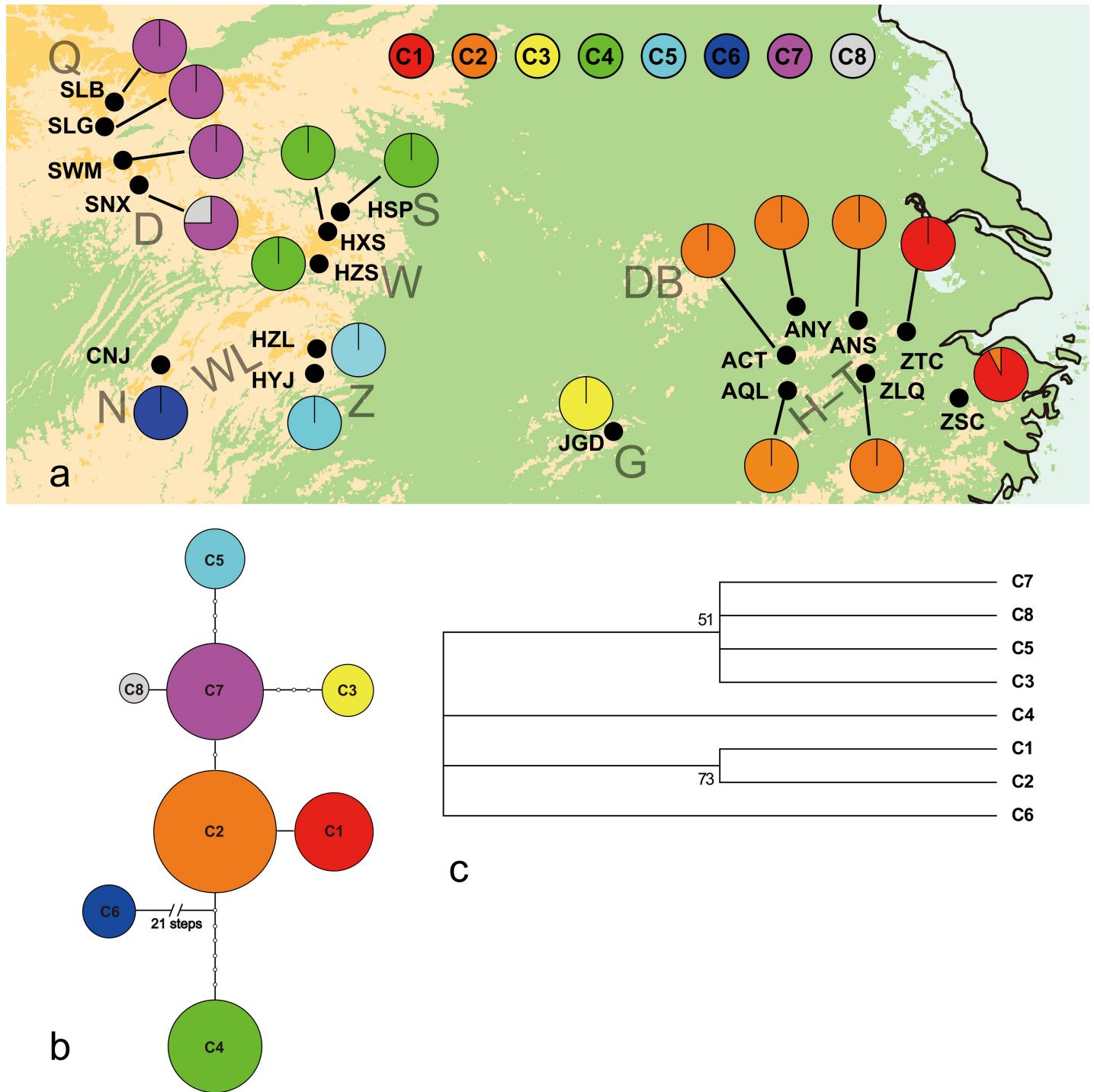

Fig 1. **(a) Distribution of 8 Chlorotypes of *B. clarkeana*.** H-T, Mt. Huang and Tianmu; G, Guan Mts.; S, Shennongjia Mts.; Z, Zhangjiajie Mts.; W, Wu Mts.; WL, Wuling Mts.; N, Nan Mts.; Q, Qinling Mts.; D, Daba Mts.; DB, Mt. Dabie. **(b) The TCS-derived Network of 8 Chlorotypes of *B. clarkeana*.** Small open circles indicate mutational steps. Circle sizes represent sample sizes (n). **(c) 50% Majority-Rule Consensus Neighbor-Joining Tree Obtained via Analysis of 8 Chlorotypes of *B. clarkeana*.** Based on 1000 permutations, bootstrap values higher than 50% are indicated above branches.

with GenAlEx v.6.501 [42]. Based on the pairwise genetic distance (D_A) calculated for the populations [45], an UPGMA tree was built using the program DISPAN [46].

Contemporary and historical gene flow

The programs BayesAss (version 1.3) and Migrate (version 3.6.5) were used to estimate contemporary (over the past few generations) and historical gene flow (over a much longer period extending back approximately $4N_e$ generations), respectively [47, 48]. Based on the Bayesian method, population-specific F_{IS} and genotypic disequilibrium were detected through Markov Chain Monte Carlo (MCMC) analysis, and contemporary gene flow was measured. Each MCMC run was performed with a burn-in of 1×10^6 and a total of 2×10^6 iterations. Five replicates for each run were performed with different initial seeds. Delta values were adjusted to maximize log likelihood values and ensure a sufficient search of the parameter space [47]. The model convergence posterior distribution was tested using Tracer v. 1.4.8 [49] to determine whether each run displayed the best model fit. The mutation-scaled migration rate (M) was estimated using Migrate based on the coalescent theory. The following equation was utilized: $M = m_h / \mu$, where μ is the mutation rate per generation (3×10^{-4}) [50], and m_h is the historical migration rate. We ran three replicates with the allele model, and the static heating scheme was applied at four temperatures (1, 1.5, 3, and 1,000,000).

Population demographic changes

For cpDNA, mismatch distribution analysis (MDA) was performed with Arlequin 3.0 [39], which was used to evaluate the evidence of spatial and demographic expansion. For each MDA model, parametric bootstrapping was used to test the goodness of fit based on the sum of squared deviations (SSD) and Harpending's raggedness index (H_{Rag}) [51]. In addition, Tajima's D [52] and Fu's F_S [53] tests of selective neutrality were performed to assess potential spatial and demographic expansion. To estimate the time (T, in number of generations) of population expansion, the expansion parameter (τ), which is also derived from MDA, and the neutral mutation rate (u) for the entire sequence per generation were used in the equation T = $\tau/2u$ [51, 54]. The u value was calculated as $u = \mu kg$, where μ is the substitution rate per site per year (s/s/y, here 1.3×10^{-9}) [15, 55], k is the average sequence length of the cpDNA region in the present study, and g is the generation time in years. The value for k in our study was 1900 bp, and 5 years was used as an approximation of the generation time, g. As no reference was available to determine generation time, we used the formula $T = \alpha + [S/(1-S)]$ to calculate the generation time T (g), where α is the maturation age and S is the survival rate [56]. Based on observations, we set the maturation age and seed survival rate to 3 years and 60%, respectively (Professor Shoubiao Zhou, Anhui Normal University, personal communication). Therefore, a generation time (T) of approximately 5 years was used in this study. TCS 1.21 [57] was employed to construct haplotype networks with the connection limit set to 95%.

Based on EST-SSRs, recent bottlenecks (past $2N_e$-$4N_e$ generations) and long-term demographic changes were detected with BOTTLENECK v. 1.2.02 and MSVAR v 1.3 [58, 59]. In the present study, two complementary methods were applied to detect genetic bottlenecks using the program BOTTLENECK. First, a mode-shift test was employed to detect the shifted model characteristics of bottlenecked populations. For non-bottlenecked populations, a large proportion of alleles were present at a low frequency, and the allele frequency curve followed a normal L-shaped distribution. When there are fewer (<10%) alleles with a low frequency than alleles with an intermediate frequency in bottlenecked populations, the normal L-shaped frequency distribution curve will shift [60]. Second, we used Wilcoxon's sign rank test. Three mutation models (infinite allele model, IAM; two-phase mutation model, TPM; stepwise mutation

model, SSM) were fitted to detect excessive heterozygosity, and a *P*-value < 0.05 using the Wilcoxon signed rank test was considered to represent evidence of a bottleneck. MSVAR was used to detect long-term changes in the effective population size and the timing of the changes. The lineal model was employed to evaluate population size changes [61], recording every 90,000 steps, with 30,000 times recorded in total. The posterior distributions of four parameters of population dynamics (ancestral effective size (N_0), current effective population size (N_1), average mutation rate across loci (μ), and time (years) since the onset of decline or expansion (T)) were estimated using Tracer v. 1.4.8 [49].

Results

CpDNA diversity and phylogeography

By concatenating the three aligned IGS regions (*psb*A-*trn*H, 325 bp; *rps*12-*rpl*20, 739 bp; *trn*L-*trn*F, 836 bp), we obtained 1900 bp of cpDNA sequence from 222 individuals. A total of 21 sites with insertion-deletions (indels) and 35 polymorphic sites (substitutions) were observed (see S1 Table). Mutations were found to be largely distributed in *psb*A-*trn*H (22 sites), mainly consisting of sequence differences between CNJ and the other populations. Only substitutions were used in the ensuing data analysis, whereas indels were excluded. In total, 8 chlorotypes (haplotypes C1-C8) were identified (see Table 1, Fig 1). In addition, multiple chlorotypes were found for only two populations (ZSC, SNX); the other 16 populations were monomorphic. Furthermore, no chlorotypes were shared among regions, even though shared chlorotypes were common among the populations in each region. Accordingly, the species exhibited high haplotype diversity and nucleotide diversity ($h_T = 0.82$, $\pi_T \times 10^3 = 3.12$) at the species level but low diversity within each population ($h_S = 0.03$, $\pi_S \times 10^3 = 0.17$). Only the eastern (Huang, Tianmu and Guan Mts.) and northwest (Qinling and Daba Mts.) regions displayed haplotype and nucleotide diversity ($h = 0.02$, $\pi_S \times 10^3 = 0.01$; $h = 0.10$, $\pi_S \times 10^3 = 0.06$), and the other two regions (central and northwest regions) showed no diversity ($h = 0$, $\pi_S \times 10^3 = 0$) (see details in S2 Table). Genetic structure was further revealed by AMOVA (see Table 2); most of the variation existed among groups (95.89%, $\Phi_{ST} = 0.994$), and significant phylogeographic structure was found ($N_{ST} = 0.995 > G_{ST} = 0.964$, both $P < 0.05$). A moderately strong relationship between the genetic and geographic distances between populations was revealed by Mantel tests ($r_M = 0.54$, $P < 0.01$). In the unrooted network (see Fig 1B), chlorotypes C2 (from Mt. Huang-Tianmu) and C7 (from Mt. Qinling-Daba) were located in the interior nodes (ancestral chlorotypes) of the network and differed by only two mutational steps, despite the great physical distances involved. In contrast, chlorotypes C4 and C5, from the adjacent Mt. Shennongjia-Wu and Mt. Zhangjiajie regions, differed by 13 mutational steps, indicating that these two regions were isolated and that their populations had evolved independently for a long time. Furthermore, chlorotypes C4 and C6 (from population CNJ, Mt. Nan) were distant from the other chlorotypes, especially C6, which exhibited more than 22 mutational steps differing from the other chlorotypes, indicating that it had been isolated from the other chlorotypes and had evolved into a distinct type over time. Chlorotypes C1, C3, C5, and C8 were subclades (tip chlorotypes) associated with C2 and C7, but separated from them by one to four mutational steps. The NJ tree (see Fig 1C) showed a similar structure to the network. Chlorotypes C6 and C4 each formed a distinct branch that was far from the other 6 chlorotypes. Chlorotypes C3 and C5 and chlorotypes from Mt. Qinling-Daba (C7, C8), which showed a close relationship, formed a large clade.

ITS diversity and phylogeography

ITS region sequences were aligned with a length of 663 bp. In total, 18 sites contained indels, and 22 showed nucleotide substitutions (see S3 Table). Only these substitutions were

Table 2. Variation percentage and expansion tests based on cpDNA sequences of *B. clarkeana*.

Groups[a]	Size	Percentage of variation			Φ_{ST}	G_{ST}(se)	N_{ST}(se)	Expansion parameter (τ)[b]	Expansion Time (t)	SSD (P)	H_{Rag} (P)	Tajima's D (P)	Fu's F_S (P)
		Among groups	Among populations	Within populations									
H-T	89	19.57	75.94	4.49	0.955	0.952 (0.036)	0.952 (0.036)	0 (0–0.016)	-	0.334 (0)	0.684 (0.940)	1.254 (0.913)	3.709 (0.923)
H-T-G	101	89.98	9.59	0.43	0.996	0.968 (0.029)	0.986 (0.015)	0 (0–0.646)	-	0.406 (0)	0.459 (0.930)	0.242 (0.649)	8.022 (0.984)
Z-W-S	60	100	0	0	1	1 (NC)	1 (NC)	3 (0–3.682)	-	0.261 (0.010)	0.701 (0.220)	2.851 (0.998)	21.511 (1.0)
WL (Z, N)	31	100	0	0	1	1 (NC)	1 (NC)	0 (0–0)	-	0.481 (0)	0.741 (1.0)	3.405 (1.0)	30.223 (1.0)
W-S-Q-D	90	98.56	0.29	1.15	0.989	0.904 (0.095)	0.986 (0.015)	0 (0–0.281)	-	0.511 (0)	0.663 (0.92)	3.247 (0.999)	20.055 (0.999)
Q-D	49	-1.26	19.4	81.86	0.181	0.182 (NC)	0.182 (NC)	3 (0.418–3.00)	121 457 (16 923–121 457)	0.02 (0.040)	0.807 (0.840)	-0.644 (0.238)	1.226 (0.581)
Total	222	95.89	3.49	0.62	0.994	0.964 (0.027)	0.995 (0.004)	39.324 (0–210.824)	-	0.061 (0.210)	0.124 (0.010)	0.029 (0.600)	22.282 (0.998)

Notes

[a]: See Table 1 for mountain codes.

[b]: 97.5% CI. N_{ST}, differentiation index of ordered haplotypes; G_{ST}, differentiation index of unordered haplotypes; *SSD*, the sum of squared deviations; *HRag*, Harpending's raggedness index.

NC, not calculated.

employed as polymorphic sites in the following data analysis; indels were excluded. Altogether, 14 ribotypes (haplotypes, R1-R14) were revealed by these polymorphic sites. Similar to the results of the cpDNA analysis, no ribotypes were shared between regions, only between populations within a single region (see Table 1, Fig 2A). These findings demonstrate that different populations within a region have commonly experienced genetic exchange of cpDNA and ITS sequences. Five populations (ZSC, ANY, HSP, HZL, SNX) contained multiple ribotypes, whereas the other 13 populations contained one fixed ribotype. For the populations with multiple ribotypes, population haplotype diversity and species nucleotide diversity ($h_T = 0.88$, $\pi_T \times 10^3 = 6.39$) in the ITS region were higher than in cpDNA. Among regions, the eastern and central regions showed more fixed ribotypes compared with the results for cpDNA. However, in contrast to the cpDNA results, the central region showed the highest haplotype diversity and nucleotide diversity ($h = 0.25$, $\pi \times 10^3 = 1.01$) among the regions (see S2 Table). Similar to the cpDNA results, the southwest region exhibited no diversity in ITS sequences, largely because only one population (CNJ) was sampled in this region. Hierarchical AMOVA revealed 75.08% ($\Phi_{ST} = 0.945$) variation between groups, indicating obvious phylogeographic structure ($N_{ST} = 0.931 > G_{ST} = 0.862$, $P = 0.059/0.037$) (see S4 Table). In contrast to the network of cpDNA chlorotypes, ribotype R1 (Mt. Huang-Tianmu) was the sole interior ribotype of the ITS network, and most ribotypes were clustered in their regions (see Fig 2B). Excluding ribotypes R6, R7, R8 (Mt. Shennongjia-Wu) and R12 (Mt. Nan), which differed from R1 by 8, 7, 6 and 6 steps, the other 10 ribotypes formed subclades (tip ribotypes) differing from R1 by one to four mutational steps. Consistent with the cpDNA results, R13 and R14 (Mt. Qinling-Daba) differed from R1 by only 2 and 1 mutational steps (Mt. Huang-Tianmu), respectively, indicating a close genetic relationship between these two regions. In contrast, the ribotypes of the Mt. Shennongjia-Wu (R6, R7 and R8) and Mt. Zhangjiajie (R9, R10 and R11) regions differed by more than 7 mutational steps, indicating long-standing isolation between these two regions.
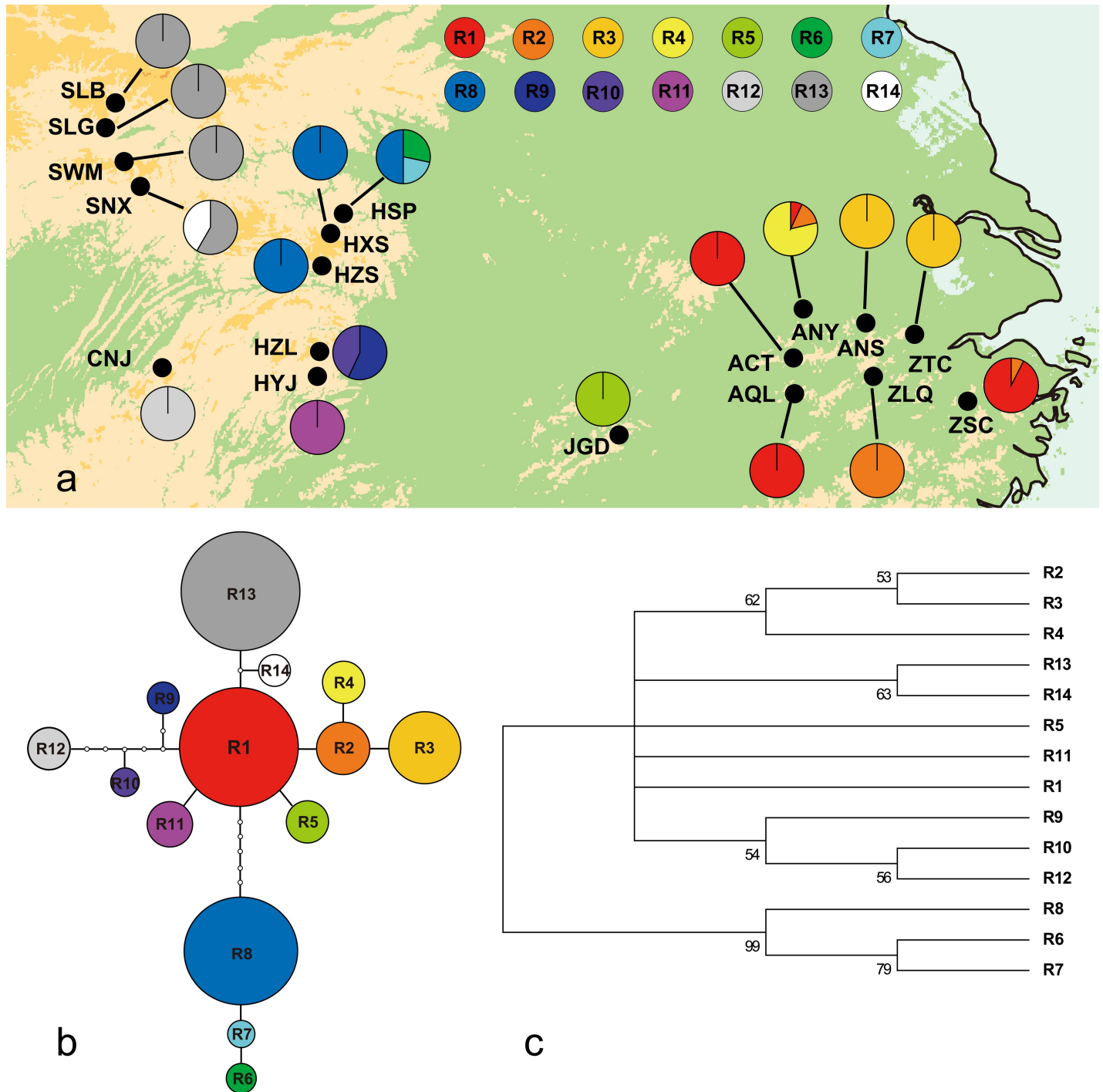
Fig 2. (a) Distribution of 14 Ribotypes in *B. clarkeana*. (b) TCS-derived Network of 14 Ribotypes of *B. clarkeana*. Small open circles indicate mutational steps. Circle sizes represent sample sizes (n). **(c) 50% Majority-Rule Consensus Neighbor-Joining Tree Obtained via the Analysis of 14 Ribotypes of *B. clarkeana*.** Based on 1000 permutations, bootstrap values higher than 50% are indicated above branches.

The NJ tree showed a similar structure to the ITS network (see details in Fig 2C). As a result, R6, R7 and R8 (Mt. Shennongjia-Wu) were grouped into one clade, which was distant from another large clade that contained the remaining 11 ribotypes. In contrast to the cpDNA

results, the CNJ population did not comprise the branch with the greatest genetic distance in this analysis.

Genetic diversity and structure based on EST-SSRs

Among the 394 individuals of *B. clarkeana* surveyed at 15 EST-SSR loci, 76 alleles were obtained, ranging from 2 to 9 per locus. The mean N_A, PIC, and H_O and H_E values were 1.34, 0.07, 0.06 and 0.08 (see Table 1), respectively, indicating very low genetic diversity in these populations. Similar to the diversity results obtained using ITS data, the central region (Wu, Shennongjia and Zhangjiajie Mts.) exhibited the highest diversity values (N_A = 1.468, PIC = 0.1, H_O = 0.096, and H_E = 0.116) among all regions, followed by the eastern region (N_A = 1.392, PIC = 0.076, H_O = 0.068, and H_E = 0.091). The northwest region (Qinling and Daba Mts.) presented the lowest diversity (N_A = 1.08, PIC = 0.013, H_O = 0.014, and H_E = 0.015) among the regions (see S2 Table for details). In addition, most of the variation existed between groups (55.23%, Φ_{ST} = 0.844), and the variation between populations within groups was higher (29.19%) than the values obtained for cpDNA and ITS. The inbreeding coefficient (F_{IS} = 0.22) was significantly positive in 11 populations, with only 4 populations (ANY, JGD, HSP and SLB) presenting a negative value, indicating common inbreeding between populations of *B. clarkeana*. The genetic differentiation between populations and regions was relatively high, and most pairwise F_{ST} values for the 18 populations were significantly greater than 0.5, except for those among the 3 populations of Mt. Qinling-Daba (SWM and SLB, 0.199; SWM and SLG, 0.007; SLG and SLB, 0.041). Additionally, the population CNJ exhibited the highest mean pairwise F_{ST} value (0.916). According to PCoA, all populations grouped into their regions, which were divided into four clusters (Mt. Huang-Tianmu-Guan, Mt. Shennongjia-Wu-Zhangjiajie, Mt. Nan and Mt. Qinling-Daba). These four regions were arranged in accordance with longitude along coordinate1 (see Fig 3). Similar to the results of PCoA, all populations in the UPGMA tree grouped into geographical regions (see Fig 4), and the population relationships showed a correlation with geographic distance. The populations of the eastern region displayed great genetic distance from the populations of the southwestern or northwestern region.

Gene flow and historical demography dynamics

Based on the EST-SSR data, contemporary and historical gene flow was detected for *B. clarkeana*. Low levels of contemporary gene flow were determined through multiple runs of BayesAss. In addition, contemporary gene flow between populations was found to be extremely weak, as the contemporary migration rate (m_c) values were all less than 0.022. The historical migration rates (m_h) calculated by Migrate ranged from 0.008 (from CNJ to SWM) to 0.248 (from SLG to SLB) (see S5 Table), with a mean value of 0.113.

Based on the results of cpDNA analysis, a unimodal mismatch distribution with non-significant *SSD* and H_{Rag} values was fitted to support the expected expansion of the distribution. In this study, Mt. Qinling-Daba was the only region satisfying the following criteria: non-significant *SSD* and H_{Rag} values (0.02, P = 0.04; 0.807, P = 0.84) and a significantly negative Tajima's D (-0.644, P = 0.238). Based on the τ-value (3), the expansion of the Mt. Qinling-Daba region was dated to approx. 121,457 yr BP (see Table 2, S1 Fig). Other than the regional expansion observed at Mt. Qinling-Daba, no major expansion or contraction of the population size of *B. clarkeana* was detected.

According to the MSVAR results, the mean effective population size of all populations of *B. clarkeana* has experienced a dramatic reduction (see Fig 5). The mean ancestral population size was larger in the eastern region than in the other regions. Population SLG exhibited the
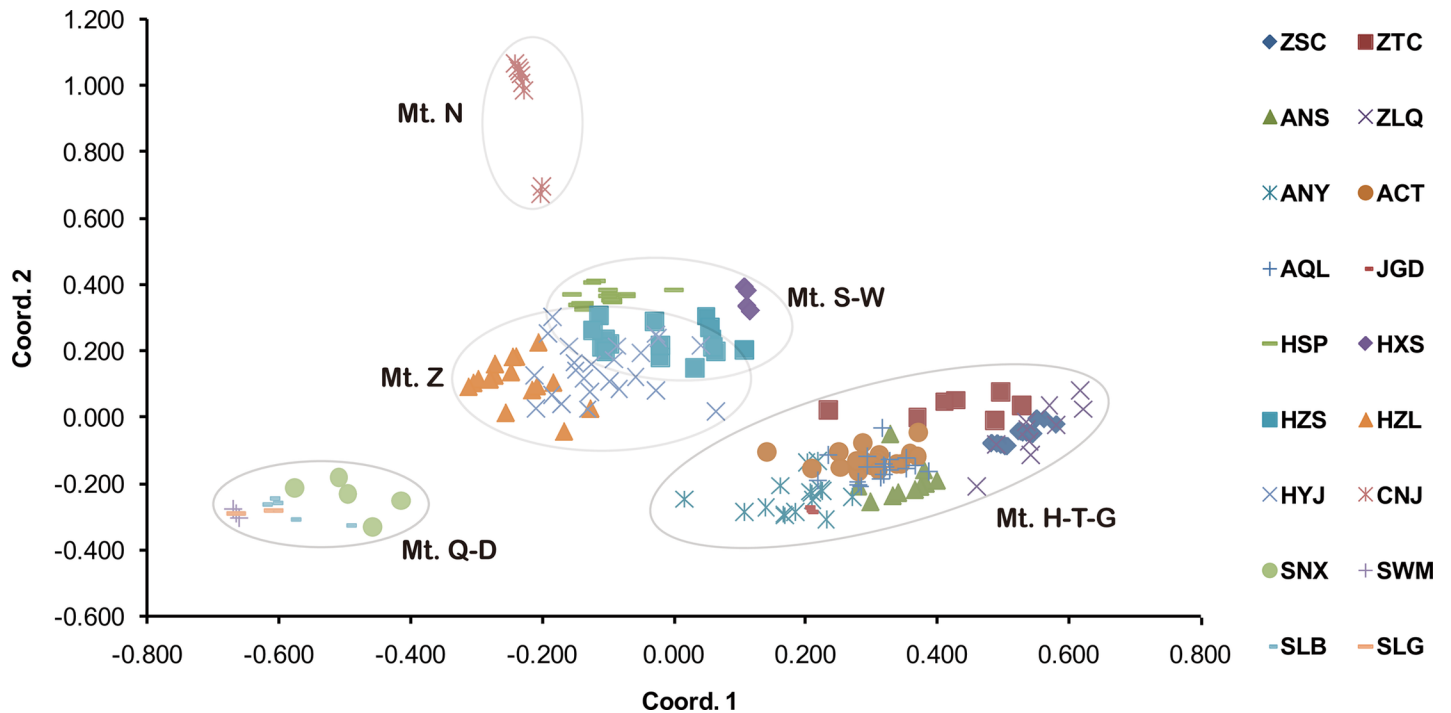
Fig 3. PCoA of 18 populations of *B. clarkeana*. See Table 1 for mountain codes.

largest current population size, with all other populations presenting a similar size. Most of the populations started to decline approx. 1,000–10,000 yr ago, except for populations SLG (~ 100,000 yr ago) and HYJ (< 1,000 yr ago). In addition, based on integrating the results regarding recent changes in effective population size tested in BOTTLENECK, seven populations showed strong evidence of recent bottlenecks. The bottlenecks mainly occurred in the eastern region, and the effect was particularly notable among the 7 populations of Mt. Huang-Tianmu, where 5 populations suffered bottleneck effects (see S6 Table). Two populations (HSP and HZL) in the central region also exhibited evidence of bottlenecks, and no bottlenecks were detected in the western region populations (Mt. Nan and Mt. Qinling-Daba). A clear progressive decrease in population bottlenecks from eastern to western regions was observed.

Discussion

Population genetic diversity

The population genetic diversity of *B. clarkeana* ($H_O = 0.06$, $H_E = 0.08$) revealed through EST-SSR analysis was significantly lower than the mean diversity values for perennial plants ($H_O = 0.63$, $H_E = 0.68$) and cross-pollinated plants ($H_O = 0.63$, $H_E = 0.65$), based on SSR data from Nybom (2004) [62]. The cpDNA and ITS data revealed the similar results. Most of the populations examined in this study showed only a single fixed haplotype, and population genetic diversity values were lowest among the studied plants in these areas [4, 7–9, 15, 63, 64]. This phenomenon of low population genetic diversity is due to several factors. First, *B. clarkeana* is a facultatively outcrossing and highly self-compatible species, and these features were reflected in the significant inbreeding coefficient [18]. Second, a substantial number of the populations may have suffered from genetic bottlenecks. Excess heterozygosity and allelic deficiency can also lead to low population genetic diversity. In addition, the habitat restriction and
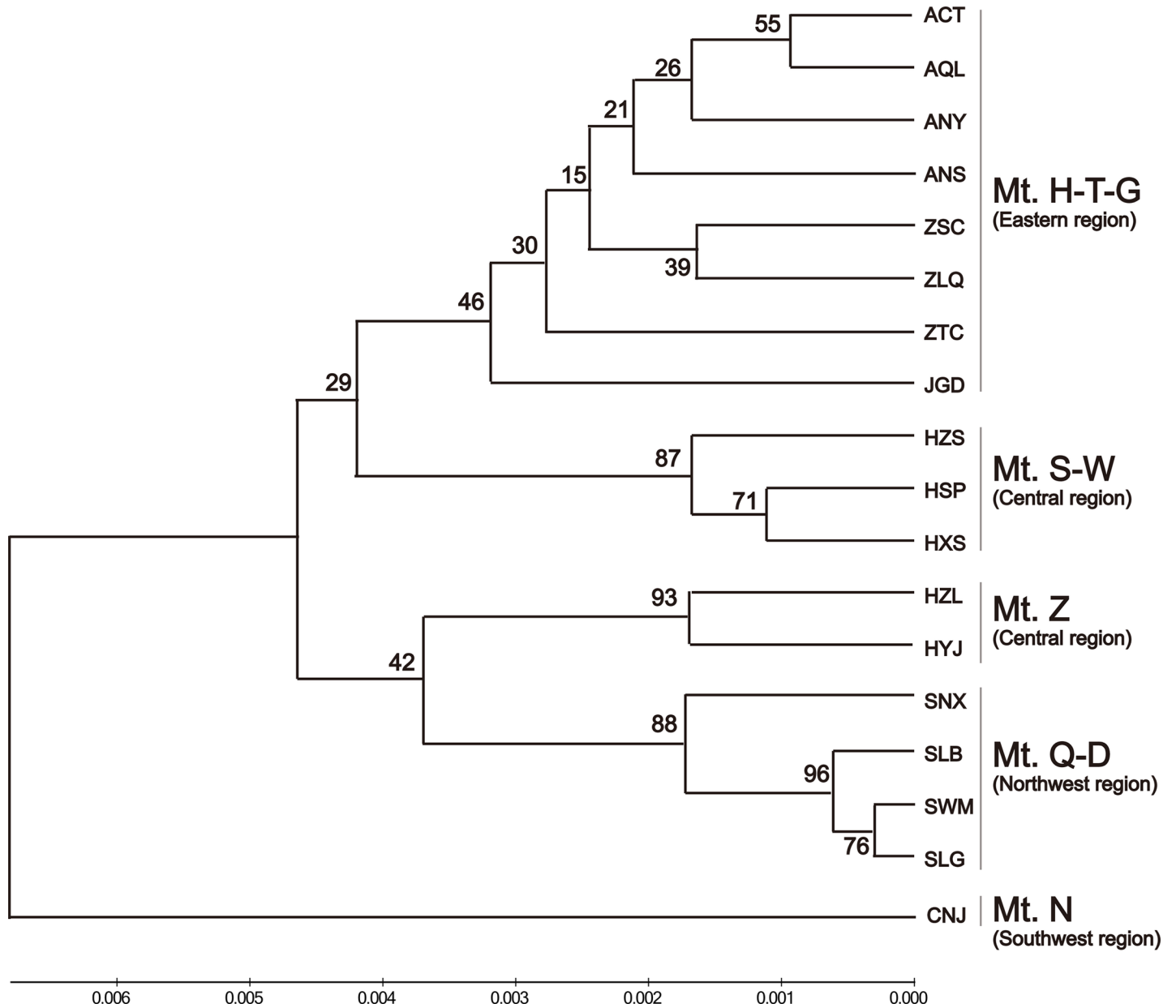
Fig 4. UPGMA dendrogram depicting Nei's (1978) genetic distances (D_A) between 18 populations of *B. clarkeana*, based on 15 EST-SSR Loci [45]. Numbers above branches indicate bootstrap values. See Table 1 for population and mountain codes.

small population sizes of this plant in its natural state are important factors leading to low genetic diversity. The short stature of *B. clarkeana* and its growth on the north side of rock out-crops (mostly limestone), requirement for scattered light and growth in the shadow of trees and shrubs are all habitat restrictions that might significantly reduce the dispersal capacity of windborne seeds and the frequency of pollination, which could lead to reproductive isolation of a population or across a region [65]. By means of orientational observations in the field, Zhang (2011) found that many populations of *B. clarkeana* exhibit a shortage of visiting polli-nators as well as a reduced frequency [21]. Moreover, field investigation have revealed a small size for most populations, which usually consist of no more than 50 individuals, while a few
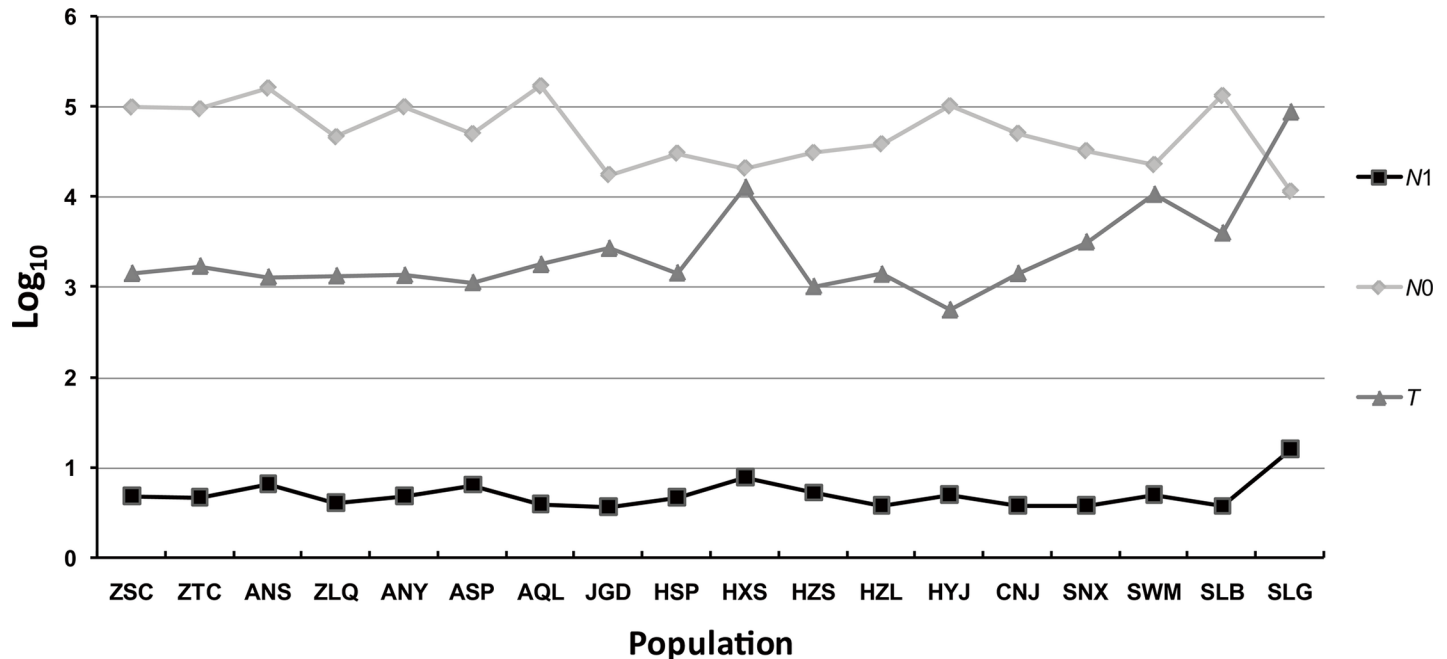
Fig 5. Effective population size and years since a change for 18 populations of *B. clarkeana*, based on the MSVAR analysis. N_1, mean current population size (over loci); N_0, mean ancestral population size; T, mean time (in years) since the population started to decline.

populations only comprise several individuals. Small plant populations are usually subject to the highest genetic risk of inbreeding and genetic drift, which can alter the genetic diversity and fitness of populations [66]. All of these characteristics decrease genetic variation within populations of *B. clarkeana*.

Genetic structure

The different genetic information and evolutionary rates between cpDNA and nuclear DNA (nDNA) [67, 68] and, thus, differences in molecular markers can result in different genetic structures [8, 12]. Accordingly, multiple molecular markers may complement each other, which would aid in understanding the genetic structure of species [69]. Based on the analysis of cpDNA, which is seed dispersed and maternally inherited, a significant barrier to seed flow was confirmed by the network and the phylogenetic tree, as we found that two adjacent regions showed haplotypes with distant relationships (Mt. Shennongjia-Wu and Mt. Zhangjiajie, Mt. Zhangjiajie and Mt. Nan). In comparison with maternal cpDNA, which exhibits biparental gene flow and a higher evolutionary rate, the nDNA EST-SSR and ITS data shared the characteristic of most populations being clustered in their regions in the generated network and phylogenetic trees. For example, the haplotype of Mt. Guan (C3) was located at a tip linked to Mt. Qinling-Daba (C7) in the cpDNA network, whereas the ribotype of Mt. Guan (R5) was located at a tip linked to Mt. Huang-Tianmu (R1). However, the PCoA and phylogenetic tree for the EST-SSR data showed a different structure from those for the cpDNA and ITS data, as the populations of Mt. Huang-Tianmu and Mt. Qinling-Daba did not exhibit a close relationship according to the EST-SSRs, which are located within expressed genes and may have undergone long-term adaptive evolution in response to the environment. The particular location in the genome as well as any transcribed products can affect the mutation patterns of SSRs [70, 71]. Due to the influence of the monsoonal wind system, precipitation and humidity on the

Chinese mainland decrease longitudinally from east to west [72, 73]. These climatic variations lead to differences in gene expression in different areas; for example, the population dehydration tolerance of *B. clarkeana* has been confirmed to increase gradually from east to west [33]. Therefore, the expressed genes containing SSRs, which may affect the EST-SSR mutation rate, might have led to the different clustering results when comparing EST-SSR with cpDNA and ITS. Overall, synthesis of the results of detection of the three markers confirmed low genetic variation within a population or region, but high genetic variation between populations from different regions [66, 74–77]. As a result, the pairwise F_{ST} values based on the EST-SSR data were significantly higher (most greater than 0.5), and the differentiation indices G_{ST} and N_{ST} (cpDNA, 0.964, 0.995; ITS, 0.862, 0.931) were also significantly higher than the mean values (maternally inherited markers, $G_{ST} = 0.655$, $N_{ST} = 0.69$; biparental, $G_{ST} = 0.163$; $N_{ST} = \sim 0.23$) between plants [78].

Glacial refugia and colonization

In this study, based on the TCS networks of cpDNA, haplotypes C2 of Mt. Huang-Tianmu and C7 of Mt. Qinling-Daba were both considered to be ancestral for *B. clarkeana*. In addition, high cpDNA haplotype and nucleotide diversity is a feature of refugia [5]. In this study, high haplotype and nucleotide diversities were only found in two populations: ZSC (Mt. Huang-Tianmu) and SNX (Mt. Qinling-Daba). These results indicate that Mt. Qinling-Daba and Mt. Huang-Tianmu represent potential refuge areas for this plant species, even though these two regions are located far apart and are near the southeastern and northwestern limits of the *B. clarkeana* distribution. The Mt. Huang-Tianmu and Mt. Qinling-Daba areas were previously shown to harbor refugia for other species during the LGM (approx. 24,000–18,000 yr BP) [5, 6, 63, 79–82].

For cpDNA, a significant and moderately strong correlation between genetic and geographic distance between the populations was revealed by Mantel tests ($r_M = 0.54$, $P < 0.01$), reflecting the 'isolation by distance' status of *B. clarkeana*. However, compared with other studies conducted in this area, the r_M value of *B. clarkeana* was even higher than that of the rare plant *Dysosma versipellis* ($r_M = 0.44$) [83] and was only lower than that of the relict *Tetracentron sinense* ($r_M = 0.62$) [9]; the r_M value for *B. clarkeana* was relatively higher. These findings indicate the occurrence of an extinction event in the Pleistocene [84]. Similar to many other plant species, the long species history of *B. clarkeana* enabled it to expand throughout the middle and lower reaches of the Yangtze River. The analyses of gene flow and effective population size variation also showed that the studied populations of *B. clarkeana* were historically large and widely distributed (significant historical gene flow and a large historical population size). However, extinction occurred due to climate change in the Pleistocene, and *B. clarkeana* was forced into two refugia (Mt. Huang-Tianmu and Mt. Qinling-Daba) by the end of Pleistocene [84–86]. According to previous studies, there is a general colonization route in this area from Mt. Huang-Tianmu to Mt. Shennongjia, passing through Mt. Dabie (see Fig 1A) [79, 87]. However, *B. clarkeana* is not currently found in Mt. Dabie, which may constitute evidence of extinction for this species. A lack of molecular variation within the regions was revealed through phylogeographic analysis (e.g., genetic variation mostly existed between regions, and no haplotypes were shared between regions), which also supports the fragmentation and extinction of *B. clarkeana* [5, 9, 84]. Because of its weak seed dispersal ability, *B. clarkeana* would not have been able to undergo long-distance expansion into suitable territories during interglacial periods, which may also have resulted in distant haplotype relationships in adjacent areas (Mt. Shennongjia-Wu and Mt. Zhangjiajie, Mt. Zhangjiajie and Mt. Nan). Therefore, no evidence that *B. clarkeana* had undergone a significant southward retreat during

glacial periods and northward migration during interglacial periods was found. Similar to some other studies [9, 88–90], no clear long colonization route was identified for *B. clarkeana*, only local expansion (from Mt. Qinling-Daba, 121,457 yr BP) occurring under suitable climatic conditions during the interglacial period (approx. 125,000–70,000 yr BP).

Human influence on effective population size reduction

According to the MSVAR results, the mean effective population size of all populations experienced a dramatic reduction, but the timing of the declines of most populations (approx. 1,000–10,000 yr ago) did not coincide with major climatic and landscape changes in this area during the Late Pleistocene (approx. 10,000–100,000 yr ago) [79]. Instead, environmental changes caused by human activities were inferred to be the main reason for the observed reduction [60, 79]. During the Yangshao Cultural period (approx. 5000–6000 yr BP), rapid agricultural growth caused serious forest degradation over a large area [91, 92], and such effects were especially prominent under the prosperous Hemudu Culture in the Yangtze River delta (5000–4500 yr BC) [93]. In this study, populations in the western region (Mt. Nan and Qinling-Daba) were not found to have been influenced by bottlenecks, although bottlenecks were detected in eastern and central regions, where intensive human activities are frequent and high-density populations exist, suggesting that bottlenecks were increased by human activities. Nevertheless, in dramatic contrast to the frequent occurrence of bottlenecks in the *B. clarkeana* populations in the lower reaches of the Yangtze River (Mt. Huang-Tianmu region), no bottlenecks were detected in *Kirengeshoma palmata* populations in this region in a previous study [60]. Although both herbs grow in wooded mountain environments, *K. palmata* lives in deep forests, whereas *B. clarkeana* grows on rocky outcrops and is more easily damaged during the clearing of land for agriculture.

Conclusions

B. clarkeana is a resurrection herb species endemic to China. In this study, the genetic structure and evolutionary history of this species were investigated based on cpDNA, ITS and EST-SSR sequence data. Similar to other plant species of Gesneriaceae, this plant is restricted to rocky-slope habitats and exhibits weak seed and pollen dispersal abilities. The genetic diversity of *B. clarkeana* was found to be exceptionally low within each population and region. More suitable habitats for *B. clarkeana* were available historically, and the species was likely to be widely distributed. Because of changes in the environment and the influence of human activity, a few populations of *B. clarkeana* have become extinct, but Mt. Qinling-Daba and Mt. Huang-Tianmu have acted as refugia, conserving ancestral haplotypes of *B. clarkeana*. Habitat loss and fragmentation are becoming increasingly serious, and we should pay attention to the fact that a small population size is becoming the norm and that the genetic risk for this plant is increasing. Moreover, some populations have suffered genetic bottlenecks, particularly in the Mt. Huang-Tianmu (Yangtze River delta) area. These analyses indicate that climate and environmental change are important for species migration and preservation. We now have a better overall understanding of the influence of habitat restriction on *B. clarkeana* and for Gesneriaceae.

Supporting information

S1 Table. CpDNA sequence polymorphisms detected in three IGS (*psb*A-*trn*H, *rpl*20-*rps*12 and *trn*L-*trn*F) regions of *B. clarkeana*, identifying 8 haplotypes.
(DOC)

S2 Table. Genetic diversity and haplotype composition of the surveyed *B. clarkeana* regions, based on cpDNA (*psb*A-*trn*H, *rps*12-*rpl*20, *trn*L-*trn*F), ITS and EST-SSR data. (DOC)

S3 Table. NrDNA sequence polymorphisms detected in the ITS region of *B. clarkeana*, Identifying 14 haplotypes. (DOC)

S4 Table. Results of analysis of molecular variation of ITS sequences from different groups of *B. clarkeana*. (DOC)

S5 Table. Mean historical migration rates among populations of *B. clarkeana*. (DOC)

S6 Table. The result of bottleneck testing in 18 populations of *B. clarkeana*. (DOCX)

S1 Fig. Distribution curve of pairwise nucleotide differences in cpDNA sequence data in the Mt. Qinling-Daba populations of *B. clarkeana*. (DOCX)

Acknowledgments

Author Contributions

Conceptualization: Ying Wang, Shoubiao Zhou.

Data curation: Ying Wang, Kun Liu, Shoubiao Zhou.

Formal analysis: Ying Wang.

Funding acquisition: Kun Liu, Jianwen Shao.

Investigation: Ying Wang, Kun Liu, De Bi.

Methodology: Ying Wang, Jianwen Shao.

Project administration: Ying Wang, Kun Liu, Shoubiao Zhou.

Resources: Ying Wang, Kun Liu, De Bi, Jianwen Shao.

Software: Ying Wang, Jianwen Shao.

Supervision: Jianwen Shao.

Validation: Ying Wang.

Visualization: Ying Wang.

Writing – original draft: Ying Wang.

Writing – review & editing: Ying Wang, Jianwen Shao.

References

1. Ying JS. Species diversity and distribution pattern of seed plants in China. Biodiv Sci. 2001; 9: 393–398.

2. Axelrod DI. Poleward migration of early angiosperm flora. British J Clin Pharmacol. 1959; 4: 141–145.

3. Wang HW, Ge S. Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. Mol Ecol. 2006; 15: 4109–4122. https://doi.org/10.1111/j.1365-294X.2006.03086.x PMID: 17054506

4. Gao LM, Moeller M, Zhang XM, Hollingsworth ML, Liu J, Mill RR, et al. High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. Mol Ecol. 2007; 16: 4684–4698. https://doi.org/10.1111/j.1365-294X.2007.03537.x PMID: 17908214

5. Gong W, Chen C, Dobeš C, Fu CX, Koch MA. Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion. Mol Phylogenet Evol. 2008; 48: 1094–1105. https://doi.org/10.1016/j.ympev.2008.05.003 PMID: 18554931

6. Tian S, Luo LC, Ge S, Zhang ZY. Clear genetic structure of *Pinus kwangtungensis* (Pinaceae) revealed by a plastid DNA fragment with a novel minisatellite. Ann Botany. 2008; 102: 69–78.

7. Wang J, Gao PX, Kang M, Lowe AJ, Huang HW. Refugia within refugia: the case study of a canopy tree (*Eurycorymbus cavaleriei*) in subtropical China. J Biogeogr. 2009; 36: 2156–2164.

8. Qi XS, Chen C, Comes HP, Sakaguchi S, Liu YH, Tanaka N, et al. Molecular data and ecological niche modelling reveal a highly dynamic evolutionary history of the East Asian Tertiary relict *Cercidiphyllum* (Cercidiphyllaceae). New Phytol. 2012; 196: 617–630. https://doi.org/10.1111/j.1469-8137.2012.04242.x PMID: 22845876

9. Sun YX, Moore MJ, Yue LL, Feng T, Chu HJ, Chen ST, et al. Chloroplast phylogeography of the East Asian Arcto-Tertiary relict *Tetracentron sinense* (Trochodendraceae). J Biogeogr. 2014; 41: 1721–1732.

10. Kou YX, Cheng SM, Tian S, Li B, Fan DM, Chen YJ, et al. The antiquity of *Cyclocarya paliurus* (Juglandaceae) provides new insights into the evolution of relict plants in subtropical China since the late Early Miocene. J Biogeogr. 2016; 43: 351–360.

11. Ma Q, Du YJ, Chen N, Zhang LY, Li JH, Fu CX. Phylogeography of *Davidia involucrata* (Davidiaceae) inferred from cpDNA haplotypes and nSSR Data. Syst Bot. 2015; 40: 796–810.

12. Cao YN, Comes HP, Sakaguchi S, Chen LY, Qiu YX. Evolution of East Asia's Arcto-Tertiary relict *Euptelea* (Eupteleaceae) shaped by Late Neogene vicariance and Quaternary climate change. BMC Evol Biol. 2016; 16: 1–17.

13. Xie XF, Yan HF, Wang FY, Ge XJ, Hu CM, Hao G. Chloroplast DNA phylogeography of *Primula ovalifolia* in central and adjacent southwestern China: Past gradual expansion and geographical isolation. J Syst Evol. 2012; 50: 284–294.

14. Comes HP, Kadereit JW. The effect of Quaternary climatic changes on plant distribution and evolution. Trends Plant Sci. 1998; 3: 432–438.

15. Qiu YX, Guan BC, Fu CX, Comes HP. Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dysosma versipellis*. Mol Phylogenet Evol. 2009; 51: 281–293. PMID: 19405195

16. Guan BC, Fu CX, Qiu YX, Zhou SL, Comes HP. Genetic structure and breeding system of a rare understory herb, *Dysosma versipellis* (Berberidaceae), from temperate deciduous forests in China. Am J Bot. 2010; 97: 111–122. https://doi.org/10.3732/ajb.0900160 PMID: 21622372

17. Liu ZW, Zhang SQ, Zhai D, Yu F, Jin ZX, Chen WF. Molecular phylogeny of the Chinese weedy rice based on nrDNA ITS and two cpDNA sequences. Indian J Genet Plant Breed. 2016; 76: 57–63.

18. Li ZY, Wang YZ. Plants of Gesneriaceae in China. Zhengzhou: Henan Science and Technology Publishing House; 2005.

19. Wen F, Li ZD. Research advance on Gesneriaceae plant. Chin Wild Plant Res. 2006; 25: 1–6.

20. Farrant JM, Brandt W, Lindsey GG. An overview of mechanisms of desiccation tolerance in selected angiosperm resurrection plants. Plant Stress. 2007; 1: 72–84.

21. Zhang D. The study on biological characteristics of a resurrection plant Boea clarkeana. Wuhu: Anhui Normal University; 2011.

22. Xiao LH, Yang G, Zhang LC, Yang XH, Zhao S, Ji ZZ, et al. The resurrection genome of *Boea hygrometrica*: A blueprint for survival of dehydration. Proc Natl Acad Sci USA. 2015; 112: 5833–5837. https://doi.org/10.1073/pnas.1505811112 PMID: 25902549

23. Jiang GQ, Wang Z, Shang HH, Yang WL, Hu ZA, Phillips J, et al. Proteome analysis of leaves from the resurrection plant *Boea hygrometrica* in response to dehydration and rehydration. Planta. 2007; 225: 1405–1420. https://doi.org/10.1007/s00425-006-0449-z PMID: 17160389

24. Liu GY, Hu YL, Zhao F, Zhu JB, Lin ZP. Molecular cloning of BcWRKY1 transcriptional factor gene from *Boea crassifolia* Hemsl. and its preliminary functional analysis. Acta Sci Nat Univ Pekin. 2007; 43: 446–452.

25. Zhu Y, Wang B, Phillips J, Zhang ZN, Du H, Xu T, et al. Global transcriptome analysis reveals acclimation-primed processes involved in the acquisition of desiccation tolerance in *Boea hygrometrica*. Plant Cell Physiol. 2015; 56: 1429–1441. https://doi.org/10.1093/pcp/pcv059 PMID: 25907569

26. Wang XQ, Li ZY. The application of sequence analysis of rDNA fragment to the systematic study of the subfamily Cyrtandroideae (Gesneriaceae). Acta Phytotax Sin. 1998; 36: 97–105.

27. Ni XW, Huang YL, Wu L, Zhou RC, Deng SL, Wu DR, et al. Genetic diversity of the endangered Chinese endemic herb *Primulina tabacum* (Gesneriaceae) revealed by amplified fragment length polymorphism (AFLP). Genetica. 2006; 127: 177–183. https://doi.org/10.1007/s10709-005-3227-0 PMID: 16850222

28. Zimmer EA, Roalson EH, Skog LE, Boggan JK, Idnurm A. Phylogenetic relationships in the Gesnerioideae (Gesneriaceae) based on nrDNA ITS and cpDNA *trn*L-F and *trn*E-T spacer region sequences. Am J Bot. 2002; 89: 296–311. https://doi.org/10.3732/ajb.89.2.296 PMID: 21669739

29. Gechev TS, Dinakar C, Benina M, Toneva V, Bartels D. Molecular mechanisms of desiccation tolerance in resurrection plants. Cell Mol Life Sci. 2012; 69: 3175–3186. https://doi.org/10.1007/s00018-012-1088-0 PMID: 22833170

30. Chao TC, Zhou SB, Chang LL, Chen YS, Xu HJ, Zhou QK. Effects of light intensity on the leaf morphology and physiological parameters of *Boea clarkeana*. Chin J Ecol. 2013; 32: 1161–1167.

31. Li ZY. The geographical distribution of the subfamily Cyrtandroideae Endl. emend. Burtt (Gesneriaceae). Acta Phytotax Sin. 1996; 34: 341–360.

32. Wang Y, Liu K, Bi De, Zhou SB, Shao JW. Characterization of the transcriptome and EST-SSR development in *Boea clarkeana*, a desiccation-tolerant plant endemic to China. PeerJ. 2017; 5: e3422. https://doi.org/10.7717/peerj.3422 PMID: 28630801

33. Zhang DD, Zhou SB, Zhou H, Liu F, Yang SY, Ma ZH. Physiological response of *Boea clarkeana* to dehydration and rehydration. Chin J Ecol. 2016; 35: 72–78.

34. Taberlet P, Gielly L, Pautou G, Bouvet J. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol. 1991; 17: 1105–1109. PMID: 1932684

35. Hamilton MB. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol Ecol. 1999; 8: 521–523. PMID: 10199016

36. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30: 2725–2729. https://doi.org/10.1093/molbev/mst197 PMID: 24132122

37. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25: 1451–1452. https://doi.org/10.1093/bioinformatics/btp187 PMID: 19346325

38. Pons O, Petit RJ. Measwring and testing genetic differentiation with ordered versus unordered alleles. Genetics. 1996; 144: 1237–1245. PMID: 8913764

39. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online. 2005; 1: 47–50.

40. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res. 1967; 27: 209–220. PMID: 6018555

41. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: a workbench to detect molecular adaptation based on a F_{st}-outlier method. BMC Bioinform. 2008; 9: 323.

42. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 2012; 28: 2537–2539. https://doi.org/10.1093/bioinformatics/bts460 PMID: 22820204

43. Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 2005; 21: 2128–2129. https://doi.org/10.1093/bioinformatics/bti282 PMID: 15705655

44. Weir BS, Cockerham CC. Estimating *F*-statistics for the analysis of population structure. Evolution. 1984; 38: 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x PMID: 28563791

45. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 1978; 89: 583–590. PMID: 17248844

46. Ota T. Dispan version 1.1: genetic distance and phylogenetic analysis. University Park, Pennsylvania: Pennsylvania State Univ. 1993.

47. Wilson GA, Rannala B. Bayesian inference of recent migration rates using multilocus genotypes. Genetics. 2003; 163: 1177–1191. PMID: 12663554

48. Beerli P, Felsenstein J. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc Natl Acad Sci USA. 2001; 98: 4563–4568. https://doi.org/10.1073/pnas.081068098 PMID: 11287657

49. Rambaut A, Drummond AJ. Tracer v1.4. Encycl Atmos Sci. 2007; 141: 2297–2305.

50. Udupa S, Baum M. High mutation rate and mutational bias at (TAA)$_n$ microsatellite loci in chickpea (*Cicer arietinum* L.). Mol Genet Genom. 2001; 265: 1097–1103.

51. Rogers AR, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol. 1992; 9: 552–569. https://doi.org/10.1093/oxfordjournals.molbev.a040727 PMID: 1316531

52. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989; 123: 585–595. PMID: 2513255

53. Fu YX. New statistical tests of neutrality for DNA samples from a population. Genetics. 1996; 143: 557–570. PMID: 8722804

54. Rogers AR. Genetic evidence for a Pleistocene population explosion. Evolution. 1995; 49: 608–615. https://doi.org/10.1111/j.1558-5646.1995.tb02297.x PMID: 28565146

55. Graur D, Li WH. Fundamentals of molecular evolution. Sunderland, MA: Sinauer Associates; 1999.

56. Lande R, Engen S, Saether BE. Stochastic population dynamics in ecology and conservation. Online: Oxford University Press; 2003.

57. Clement M, Posada D, Crandall KA. TCS: a computer program to estimate gene genealogies. Mol Ecol. 2000; 9: 1657–1659. PMID: 11050560

58. Piry S, Luikart G, Cornuet JM. Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. J Heredity. 1999; 90: 502–503.

59. Girod C, Vitalis R, Leblois R, Fréville H. Inferring population decline and expansion from microsatellite data: a simulation-based evaluation of the Msvar method. Genetics. 2011; 188:165–179. https://doi.org/10.1534/genetics.110.121764 PMID: 21385729

60. Yuan N, Sun Y, Comes HP, Fu CX, Qiu YX. Understanding population structure and historical demography in a conservation context: population genetics of the endangered *Kirengeshoma palmata* (Hydrangeaceae). Am J Bot. 2014; 101: 521–529. https://doi.org/10.3732/ajb.1400043 PMID: 24650862

61. Beaumont MA. Detecting population expansion and decline using microsatellites. Genetics. 1999; 153: 2013–2029. PMID: 10581303

62. Nybom H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol. 2004; 13: 1143–1155. https://doi.org/10.1111/j.1365-294X.2004.02141.x PMID: 15078452

63. Qiu YX, Sun Y, Zhang XP, Lee J, Fu CX, Comes HP. Molecular phylogeography of East Asian *Kirengeshoma* (Hydrangeaceae) in relation to Quaternary climate change and landbridge configurations. New Phytol. 2009; 183: 480–495. https://doi.org/10.1111/j.1469-8137.2009.02876.x PMID: 19496955

64. Huang DQ, Li QQ, Zhou CJ, Zhou SD, He XJ. Intraspecific differentiation of *Allium wallichii* (Amaryllidaceae) inferred from chloroplast DNA and internal transcribed spacer fragments. J Syst Evol. 2014; 52: 341–354.

65. Zhang XP, Li XH, Qiu YX. Genetic diversity of the endangered species *Kirengeshoma palmata* (Saxifragaceae) in China. Biochem Syst Ecol. 2006; 34: 38–47.

66. Ellstrand NC, Elam DR. Population genetic consequences of small population size: implications for plant conservation. Annu Rev Ecol Syst. 1993; 24: 217–242.

67. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA. 1987; 84: 9054–9058. PMID: 3480529

68. Hancock JM. Microsatellites and other simple sequences: Genomic context and mutational mechanisms. In: Goldstein DB, Schlötterer C, editors. Microsatellites: evolution and applications. New York: Oxford University Press; 1999.

69. Lowe A, Harris S, Ashton P. Ecological genetics: design, analysis, and application. Carlton, Australia: Blackwell Publishing company; 2004.

70. Anmarkrud JA, Kleven O, Bachmann L, Lifjeld JT. Microsatellite evolution: mutations, sequence variation, and homoplasy in the hypervariable avian microsatellite locus HrU10. Bmc Evol Biol. 2008; 8: 138. https://doi.org/10.1186/1471-2148-8-138 PMID: 18471288

71. Hawk JD, Stefanovic L, Boyer JC, Petes TD, Farber RA. Variation in efficiency of DNA mismatch repair at different sites in the yeast genome. Proc Natl Acad Sci USA. 2005; 102: 8639–8643. https://doi.org/10.1073/pnas.0503415102 PMID: 15932942

72. Zhai PM, Zhang XB, Wan H, Pan XH. Trends in total precipitation and frequency of daily precipitation extremes over China. J Climate. 2005; 18: 1096–1108.

73. Wan SM, Li AC, Clift PD, Stuut JBW. Development of the East Asian monsoon: Mineralogical and sedimentologic records in the northern South China Sea since 20Ma. Palaeogeogr Palaeoclimatol Palaeoecol. 2007; 254: 561–582.

74. Hamrick JL, Godt MJW, Brown AHD, Clegg MT, Kahler AL, Weir BS. Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS, editors. Plant Population Genetics Breeding and Genetic Resources. Sunderland, MA: Sinauer Associates; 1989. pp. 43–63.

75. Karron JD. A comparison of levels of genetic polymorphism and self-compatibility in geographically restricted and widespread plant congeners. Evol Ecol. 1987; 1: 47–58.

76. Karron JD, Falk DA, Holsinger KE. Patterns of genetic variation and breeding systems in rare plant species. In Falk DE, Holsinger KE, editors. Genetics and conservation of rare plants. New York: Oxford University Press; 1991. 87–98.

77. Crawford DJ, Stuessy TF, Haines DW, Cosner MB, O. MS, Lopez P. Allozyme diversity within and divergence among four species of *Robinsonia* (Asteraceae: Senecioneae), a genus endemic to the Juan Fernandez Islands, Chile. Am J Bot. 1992; 79: 962–966.

78. Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. Invited Review: Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol Ecol. 2005; 14: 689–701. https://doi.org/10.1111/j.1365-294X.2004.02410.x PMID: 15723661

79. Qiu YX, Fu CX, Comes HP. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol Phylogenet Evol. 2011; 59: 225–244. https://doi.org/10.1016/j.ympev.2011.01.012 PMID: 21292014

80. Tian B, Liu RR, Wang LY, Qiang Q, Chen KM, Liu JQ. Phylogeographic analyses suggest that a deciduous species (*Ostryopsis davidiana* Decne., Betulaceae) survived in northern China during the Last Glacial Maximum. J Biogeogr. 2009; 36: 2148–2155.

81. Bai WN, Liao WJ, Zhang DY. Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. New Phytol. 2010; 188: 892–901. https://doi.org/10.1111/j.1469-8137.2010.03407.x PMID: 20723077

82. Li EX, Yi S, Qiu YX, Guo JT, Comes HP, Fu CX. Phylogeography of two East Asian species in *Croomia* (Stemonaceae) inferred from chloroplast DNA and ISSR fingerprinting variation. Mol Phylogenet Evol. 2008; 49: 702–714. https://doi.org/10.1016/j.ympev.2008.09.012 PMID: 18849001

83. Qiu YX, Li JH, Liu HL, Chen YY, Fu CX. Population structure and genetic diversity of *Dysosma versipellis* (Berberidaceae), a rare endemic from China. Biochem Syst Ecol. 2006; 34: 745–752.

84. Koch M, Bernhardt KG. Comparative biogeography of the cytotypes of annual *Microthlaspi perfoliatum* (Brassicaceae) in Europe using isozymes and cpDNA data: refugia, diversity centers, and postglacial colonization. Am J Bot. 2004; 91: 115–124. https://doi.org/10.3732/ajb.91.1.115 PMID: 21653368

85. Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M. Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. Mol Ecol. 2003; 12: 201–212. PMID: 12492888

86. Trewick SA, Morganrichards M, Russell SJ, Henderson S, Rumsey FJ, Pintér I, et al. Polyploidy, phylogeography and Pleistocene refugia of the rockfern *Asplenium ceterach*: evidence from chloroplast DNA. Mol Ecol. 2002; 11: 2003–2012. PMID: 12296944

87. Zhang YH. Phylogeography and landscape genetics of Emmenopterys henryi (Rubiaceae), a Tertiary relict species across China. Hangzhou: Zhejiang University; 2016.

88. Yu H, Nason JD. Nuclear and chloroplast DNA phylogeography of *Ficus hirta*: obligate pollination mutualism and constraints on range expansion in response to climate change. New Phytol. 2013; 197: 276–289. https://doi.org/10.1111/j.1469-8137.2012.04383.x PMID: 23127195

89. Zhang JJ, Li ZZ, Fritsch PW, Tian H, Yang AH, Yao XH. Phylogeography and genetic structure of a Tertiary relict tree species, *Tapiscia sinensis* (Tapisciaceae): implications for conservation. Ann Botany. 2015; 116: 727–737.

90. Wang YH, Jiang WM, Comes HP, Hu FS, Qiu YX, Fu CX. Molecular phylogeography and ecological niche modelling of a widespread herbaceous climber, *Tetrastigma hemsleyanum* (Vitaceae): insights into Plio-Pleistocene range dynamics of evergreen forest in subtropical China. New Phytol. 2015; 206: 852–867. https://doi.org/10.1111/nph.13261 PMID: 25639152

91. Yi S, Saito Y, Zhao QH, Wang PX. Vegetation and climate changes in the Changjiang (Yangtze River) Delta, China, during the past 13,000 years inferred from pollen records. Quat Sci Rev. 2003; 22: 1501–1519.

92. Ren GY, Beug HJ. Mapping Holocene pollen data and vegetation of China. Quat Sci Rev. 2002; 21: 1395–1422.

93. Cai BQ. A query on Si agriculture in Hemudu culture. J Xiamen Univ. 2006; 1: 49–55.