Strong population bottleneck and repeated demographic expansions of Populus adenopoda (Salicaceae) in subtropical China

Original Article

Strong population bottleneck and repeated demographic expansions of *Populus adenopoda* (Salicaceae) in subtropical China

Liqiang Fan[1,2,†], Honglei Zheng[1,†], Richard I. Milne[3], Lei Zhang[1] and Kangshan Mao[1,*]

[1]*Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, Sichuan, P. R. China*

[2]*State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang 455000, Henan, P. R. China*

[3]*Institute of Molecular Plant Sciences, University of Edinburgh, Edinburgh EH9 3JH, UK*

[†]These authors contributed equally to this work.

Running title: A dynamic Quaternary evolutionary history of the Chinese aspen

*For correspondence. E-mail: maokangshan@scu.edu.cn or maokangshan@163.com

Received date: 18 August 2017 Returned for Revision date: 25 September 2017

Editorial Decision Date: 29 November 2017

ABSTRACT

- **Background and Aims** Glacial refugia and inter-/postglacial recolonization routes during the Quaternary of tree species in Europe and North America are

well understood, but far less is known about those of tree species in subtropical eastern Asia. Thus, we have examined the phylogeographic history of *Populus adenopoda* (Salicaceae), one of the few poplars that naturally occur in this subtropical area.

- **Methods** Genetic variations across the range of the species in subtropical China were surveyed using 10 nuclear microsatellite loci and four chloroplast fragments (*mat*K, *trn*G-*psb*K, *psb*K-*psb*I and *ndh*C-*trn*V). Coalescent-based analyses were used to test demographic and migration hypotheses. In addition, species distribution models (SDMs) were constructed to infer past, present and future potential distributions of the species.
- **Key Results** Thirteen chloroplast haplotypes were detected, and haplotype-rich populations were found in central and southern parts of the species' range. STRUCTURE analyses of nuclear microsatellite loci suggest obvious lineage admixture, especially in peripheral and northern populations. DIYABC analysis suggests that the species might have experienced two independent rounds of demographic expansions and a strong bottleneck in the late Quaternary. SDMs indicate that the species' range contracted during the last glacial maximum (LGM), and contracted northward but expanded eastward during the last interglacial (LIG).
- **Conclusions** Chloroplast data and SDMs suggest that *P. adenopoda* might have survived in multiple glacial refugia in central and southern parts of its

range during the LGM. Populations of the Yunnan-Guizhou Plateau in the southern part have high cpDNA diversity, but may have contributed little to the post-glacial recolonization of northern and eastern parts. The three major demographic events inferred by DIYABC coincide with the initiation of the LIG, start of the LGM and end of the LGM, respectively. The species may have experienced multiple rounds of range contraction during glacial periods and range expansion during interglacial periods. Our study corroborates the importance of combining multiple lines of evidence when reconstructing Quaternary population evolutionary histories.

Key words: Microsatellite, multiple loci, species distribution modeling, demographic history, coalescent tests, Quaternary.

INTRODUCTION

The impacts of Quaternary climate oscillations, during which glacial and interglacial cycles alternated, on organisms' distributions has attracted tremendous biogeographical attention (Bennett *et al.*, 1991; Hewitt, 2000; Tzedakis *et al.*, 2013). In recent decades, various approaches have been developed and refined to elucidate climate refugia, postglacial recolonization routes and associated demographic changes (Carstens and Richards, 2007; Avise, 2009; Gavin *et al*., 2014). One is construction of species distribution models (SDMs), for predicting past, present and future potential geographical distributions of species based on their ecological niches (Carstens and

Richards, 2007; Elith and Leathwick, 2009). Another is model-based statistical phylogeographic inference (Knowles and Maddison, 2002; Carstens and Richards, 2007; Hickerson *et al.*, 2010) using methods such as approximate Bayesian computations (ABC) to test series of complex biogeographic scenarios under a coalescent framework (Beaumont, 2010). These approaches, together with the increasing abundance of fossil records (Tzedakis *et al.*, 2013), have enabled increasingly sophisticated reconstructions and compilations of Quaternary histories of species based on multiple lines of evidence (Carstens and Richards, 2007; Hickerson *et al.*, 2010; Gavin *et al.*, 2014). Nevertheless, most studies concerning effects of climate oscillations have focused on species in Europe and North America (Hewitt, 2000; Shafer *et al.*, 2010; Tzedakis *et al.*, 2013), while Quaternary range dynamics and genetic consequences for species in other regions (for example, subtropical China) remain less well understood (e.g. Qiu *et al.*, 2011; Tian *et al.*, 2015; Kou *et al.*, 2016; Zhang *et al.*, 2016).

The subtropical China region usually refers to the hilly mid-elevation (Wu, 1980; Wu *et al.*, 1987) areas in China bordered by the Qingling Mountains-Huai River line (ca. 34° N) in the north, the tropics (ca. 22° N) in the south, the Qinghai-Tibetan Plateau in the west, and the coastline in the east. This geographically and climatically heterogeneous region covers around 2.4 million km^2, more than 25% of the area of China (Wang *et al.*, 2009). It incorporates large parts of the three major plant diversity hotspots in China—the Hengduan mountains, Central China and Lingnan areas (Ying,

2000)—plus various finer-scale biodiversity hotspots in the west (López-Pujol *et al.*, 2011). The current flora in this region is dominated by warm-temperate to subtropical evergreen forests, intermingled with isolated patches of warm-temperate to subtropical deciduous forests. Mean annual precipitation amounts to ca. 1814 mm, and the mean annual air temperature is ca. 11.5°C (Zhou *et al.*, 2015). During the Last Glacial Maximum (LGM), the climate in this region was ca. 4-6°C cooler (Zheng *et al.*, 2003) and precipitation was ca. 400-600 mm/yr lower (Qiu *et al.*, 2011; Lu *et al.*, 2013) than currently, partly due to a stronger winter monsoon. Reconstructed paleo-biome maps based on pollen records from the LGM indicate that plant species in the subtropical China region may have experienced latitudinal and altitudinal migrations, population extinctions, and range fragmentation and/or expansion due to Quaternary climate oscillations (Yu *et al.*, 2000; Harrison *et al.*, 2001; Qiu *et al.*, 2011).

Previous phylogeographic studies have revealed multiple glacial refugia for tree species in subtropical China, and various woody species are likely to have experienced southward glacial retreat and northward, eastward or localized postglacial range expansions (Qiu *et al*., 2011; Qi *et al*., 2012; Shi *et al*., 2014; Sun *et al*., 2014a; Gong *et al.*, 2016). However, fewer such studies have fully covered the western part of subtropical China than eastern and southern parts of the region (Peng *et al.*, 2015; Gong *et al.*, 2016). Therefore, the role of postglacial expansion in the assembly of the flora of western subtropical China remains obscure.

Populus adenopoda Maxim. (Salicaceae), the Chinese aspen, is a deciduous and wind-pollinated tree of the subtropical China region (Fang *et al.*, 1999). It fruits around April to May, and produces numerous "cotton-embraced" seeds that are well adapted to wind-aided dispersal (Fang *et al.*, 1999). It occurs from 300-2500 m a.s.l. on sunny slopes or along riversides, and is common in southwestern China, especially Guizhou, but scattered and rare in central and eastern China, with scant herbarium records. In Guizhou, this aspen is reportedly a pioneer tree species in burnt or disturbed sites, and it usually forms small pure stands or dominates mixed forest (Jiang and Tan, 2009). Given its ability to rapidly colonize vacant habitats, *P. adenopoda* might be an ideal model to track range dynamics of vegetation in subtropical China.

Thus, we have investigated the phylogeography of *P. adenopoda* by sequencing four maternally inherited chloroplast (cpDNA) markers (*mat*K, *trn*G-*psb*K, *psb*K-*psb*I and *ndh*C-*trn*V) and genotyping 10 bi-parentally inherited nuclear microsatellite (nSSR) loci across 39 populations covering the main geographic distribution of the species. We conducted various phylogeographic analyses, including coalescent-based DIYABC, and constructed SDMs to infer the species' potential distribution during the LIG (last interglacial), the LGM (last glacial maximum), the MH (middle Holocene), the present day and the future. We addressed the following three questions. Where were the glacial refugia (suitable habitats) for *P. adenopoda* during the late Quaternary, and are they consistent with previous phylogeographic studies and

paleo-biome reconstructions in this region? Did the species' range contract and expand along latitudinal gradients following Quaternary climate changes? Are there any phylogeographic and demographic indications of repeated range expansions and contractions?

MATERIALS AND METHODS

Sampling and data collection

Samples from 39 populations across the geographic distribution of *P. adenopoda* in subtropical China were collected (Fig. 1; Editorial Committee of Vegetation Map of China, 2007). Fresh leaves of 5-20 individuals per population were sampled, dried and stored with silica gel. Sampled individuals of 37 populations were separated by at least 50 m, but individuals of populations 35 and 37 were only separated by at least 20 m, because there were too few individuals with larger spacing. The altitude, latitude, and longitude of each sampled population were measured using an Extrex GIS monitor (Garmin, Taiwan) (Table S1).

We examined nSSR and cpDNA sequences in 415 and 371 individuals of the 39 populations, respectively (Table S1). An individual of *P. davidiana* was sampled as an outgroup, to root the cpDNA haplotype network (see below). Genomic DNA was extracted from the silica-dried leaves of each individual using the CTAB method (Doyle and Doyle, 1987). Three cpDNA fragments (*mat*K, *trn*G-*psb*K and *psb*K-*psb*I) were amplified using primers described by Schroeder *et al.* (2012) (Table S2a). In

addition, *ndh*C-*trn*V was amplified using newly designed primers (Table S2a). PCR reactions were carried out as described in Table S2a. The PCR products were checked on 1% agarose gels and sent to Tsingke Biological Technology (Beijing, China) for sequencing. Chloroplast DNA sequences have been deposited in NCBI GenBank (accession nos. KX925263-KX925275).

After a preliminary screen of 24 nSSR loci using DNA from 48 individuals, 12 nSSR loci (Table S2b; Tuskan *et al*., 2004; Ma *et al*., 2013; Jiang *et al*., 2016) were selected based on PCR success rate and polymorphism, for genotyping each individual. PCR reactions were performed as described in Table S2b. The PCR products were checked on 1% agarose gels and sent to Honor Tech (Beijing, China) for nuclear microsatellite genotyping. Nuclear microsatellite data are available at ResearchGate (http://www.researchgate.net, doi: 10.13140/RG.2.2.12841.80480).

Statistical analyses of cpDNA sequence data

The cpDNA sequences were proofread and aligned using MEGA v. 6.06 (Tamura *et al.*, 2013), then haplotypes were detected using DNASP v. 5.10 (Librado and Rozas, 2009). Insertions/deletions (Indels, excluding mononucleotide repeats) were coded by the simple gap coding method (Simmons and Ochoterena, 2000) as implemented in Gapcoder (Young and Healy, 2003), then the 0/1 characters (except '-' gaps after coding) were replaced manually by A/T to use indel information (e.g. Havrdová *et al*., 2015). A median-joining haplotype network was constructed using NETWORK v.

4.5.1 (Bandelt *et al.*, 1999) under the maximum parsimony (MP) criterion for the 39 populations. We also constructed a median-joining haplotype network for these populations together with 76 populations of the *P. davidana-rotundifolia* complex, using samples collected in a previous study (Zheng *et al*., 2017) to detect possible shared cpDNA haplotypes between *P. adenopoda* and the complex.

Analysis of molecular variance (AMOVA) was performed to partition variation within and among populations using ARLEQUIN v. 3.11 (Excoffier *et al.*, 2005). We also tested the presence of significant phylogeographic structure (the correlation between haplotypes' genetic and geographic closeness), by calculating and comparing G_{ST} and N_{ST} statistics based on 1000 permutations in PERMUT (Pons and Petit, 1996). When interpreting the results, we bore in mind that lack of significant phylogeographic structure does not necessarily mean lack of population structure.

To test for historical demographic expansions, mismatch distributions were implemented using ARLEQUIN, based on 1,000 parametric bootstrap replicates. We also conducted neutrality tests by calculating Tajima's *D* and Fu's F_S values in ARLEQUIN.

Statistical analyses of nuclear microsatellite data

Microsatellite primers were synthesized with the 5' end of the forward primer of each pair tagged with either a 6-FAM, a TAMR or a HEXA fluorescent label, for automated fragment analysis using an ABI PRISM 3100 genetic analyzer (Applied

Biosystems, Foster City, CA, USA). Then, the nuclear microsatellite genotype at each locus in each individual was determined using GeneMarker (Softgenetics, Pennsylvania, USA) and the FLEXIBIN Excel macro (Amos *et al.*, 2007). Finally, allele sizes at each locus were scored and checked for possible genotyping errors like stuttering, large allele dropouts and null alleles in CERVUS (Dakin and Avise, 2004). Two loci at which null alleles ($F_{(Null)} > 0.4$) were detected were eliminated and the remaining 10 nSSR loci were subjected to further analyses (Table S2b). *Populus adenopoda* is known to reproduce clonally through basal shoots. This trait could enable the collection of samples from identical individuals, thereby interfering with results, but we minimized this risk by collecting samples from trees separated by at least 50 m in 37 populations, and least 20 m in the other two populations. We identified individuals using the nSSR data and GIMLET v. 1.3.3 (Valière, 2002).

Five parameters of genetic variation within populations (A, mean number of alleles per locus; Ae, the effective number of alleles; Ar_5, allelic richness based on five samples; H_o, observed heterozygosity; and H_e, expected heterozygosity) were quantified using GenAlEx v. 6.0 (Peakall and Smouse, 2006) and Microsatellite Analyser (MSA) v. 4.05 (Dieringer and Schlötterer, 2003). We also analyzed partitioning of total genetic variation among groups, within populations and among populations by AMOVA, with 1000 permutations using ARLEQUIN.

We used the individual-based Bayesian clustering method implemented in STRUCTURE v. 2.3.4 (Pritchard *et al.*, 2000) to investigate population subdivision.

For each *K* value between 1 and 10, we performed 20 independent simulations with 500,000 burn-in steps followed by 1500,000 MCMC steps using the admixture model with correlated allele frequencies. The most likely numbers of clusters (*K*) were evaluated using both the likelihood estimation (Pritchard *et al.*, 2000) and Δ*K* statistic (Evanno *et al.*, 2005) method, implemented in STRUCTURE HARVESTER (Earl and von Holdt, 2012). These analyses were conducted both for the 39 sampled *P. adenopoda* populations, and for these populations together with 76 populations of the *P. davidana-rotundifolia* complex, using samples collected in our previous study (Zheng *et al*., 2017) based on 10 nSSR loci, to detect possible sharing of genetic components between *P. adenopoda* and the complex.

To elucidate the demographic history of *P. adenopoda*, we employed the approximate Bayesian computation (ABC) framework in DIYABC v. 2.0 (Cornuet *et al*., 2014) to test six plausible scenarios based on nSSR loci (Fig. 2), using parameter settings presented in Table S3. Both a direct approach and logistic regression analyses were applied to estimate the posterior probability of the best explanatory scenario and subsequent parameters of interest. The scenario with the highest posterior probability was selected, and the associated parameters were determined. Since estimates of time in DIYABC are in generation numbers, we assumed a generation time of 40 years for *P. adenopoda* to transform generation numbers into calendar years. The generation time of *Populus* was assumed to be 15 years by Levsen *et al.* (2012) and 20 to 60 years by Macaya-Sanz *et al*. (2012). We adopted the average generation time of the

latter study since we are studying *P. adenopoda*'s evolutionary history during a very long time interval, and individuals probably took longer to reach sexual maturity in the harsh environments that prevailed during at least some past periods, especially glacial periods.

Isolation by distance, BARRIER and MIGRATE analyses

Isolation-by-distance was assessed with GenAIEx, by testing the correlation between the genetic differentiation, quantified as $F_{ST}/(1–F_{ST})$, and the corresponding geographical distance (km) between the 39 populations represented in the nSSR dataset. Monmonier's maximum-difference algorithm implemented in BARRIER v. 2.2 (Manni *et al.*, 2004) was used to identify biogeographic boundaries, or areas marking the largest genetic discontinuities between population pairs. The robustness of these boundaries was assessed by applying the algorithm to 1000 Nei's distance matrices, produced by bootstrapping over loci with MSA.

Historical gene flow among different population groups (as identified by BARRIER analysis) was assessed using MIGRATE v. 3.2.6 (Beerli and Felsenstein, 1999; Beerli and Palczewski, 2010). The amount and direction of gene flow were estimated from the nSSR data by calculating the parameters N_e (effective population size) and effective number of migrants ($2N_em$, where N_e is effective population size and m is the migration rate). Using the continuous Brownian motion model, we started five independent MCMC chains, each with 5,000,000 generations and sampled

every 100 steps under a constant mutation model, discarding the first 1,000,000 records as burn-in. The mode and 95% highest posterior density were then estimated after checking for data convergence.

Species distribution modeling

Species distribution models (SDMs) provide a means to predict potential historical changes in species' distributions independently of any genetic approach. We used SDMs to determine potential distributions of *P. adenopoda*, using the maximum entropy approach as implemented in MAXENT v. 3.3.3k (Phillips *et al.*, 2006), in six periods. These were the Last Interglacial (LIG, ca. 140-120 thousand years ago [Ka]), Last Glacial Maximum (LGM, ca. 21-18 Ka), Middle Holocene (MH, ca. 6 Ka), the present-day and two future scenarios, for the years 2050 (model rcp45, average over 2041-2060, assuming a radiative forcing value of 4.5/Wm^2) and 2070 (model rcp85, average over 2061-2080, assuming a radiative forcing value of 8.5/Wm^2). The present species occurrence data for *P. adenopoda* consisted of our 39 sampling sites (Table S1, field survey dataset) and records of the species' occurrence compiled since the 1980s in the Chinese Virtual Herbarium (www.cvh.org.cn). After eliminating potentially misidentified samples and specimens with recorded sampling sites that are too vague to trace, 81 points (42 from the herbarium records and 39 from our field data) were retained and used in subsequent modelling.

We obtained the elevation variable and 19 bioclimatic variables for the six time-periods at a resolution of 2.5 arc-mins from the WorldClim database (http://www.worldclim.org; Hijmans *et al.*, 2005). To avoid model over-fitting, only the eight environmental variables with pairwise Pearson correlation coefficients less than 0.7 were retained for SDMs. Graphic displays of distributions predicted by each SDM were generated using DIVA-GIS v. 7.5 (Hijmans *et al.*, 2012).

RESULTS

Variations in the chloroplast DNA sequences

The total alignment length of the four cpDNA fragments in the 371 individuals examined was 2,078 bp. Nucleotide substitution at one site and indels at nine positions were detected (Table S4), and 13 haplotypes (H1 to H13; Table S1) were identified. In 11 populations a single haplotype was fixed. In addition, two, three, four and five haplotypes were detected in 20, five, two and one populations, respectively; and all of the eight populations with ≥3 haplotypes occupied sites south of the Yangtze River, mostly on or around the Yunnan-Guizhou (Y-G) Plateau. Private haplotypes were found in eight populations (1, 2, 6, 12, 18, 20, 22, 24), nearly all located in southern and central parts of the species' range. The haplotype frequencies for each population and their geographical distributions are shown in Table S1 and

Fig. 3. No haplotype shared by *P. adenopoda* and the *P. Davidiana-rotundifolia* complex was found (Fig. S1).

According to AMOVA, 79.07% of the cpDNA variation was within populations, and only 20.93% among populations (F_{ST} = 0.20931, P < 0.01) (Table 1). Moreover, a permutation test revealed that G_{ST} (0.328 ± 0.063) exceeded N_{ST} (0.323 ± 0.049), but not significantly (P > 0.05). Thus, we did not detect significant phylogeographic structure in the sampled populations of *P. adenopoda* (Pons and Petit, 1996).

Results of our mismatch distribution analysis based on the cpDNA sequence data are indicative of demographic expansion. Neither results of further variance analyses nor the raggedness index obtained significantly deviated from expectations based a model of sudden population expansion (Table 2). Nevertheless, Fu's and Tajima's neutrality tests respectively yielded non-significant negative F_S (-0.79739, P > 0.05) and D (-0.5768, P = 0.272) values (Table 2).

Genetic diversity, division of population groups and asymmetrical migrations between groups based on nuclear microsatellite markers

Among the 10 nSSR loci, we identified 3.5 alleles per locus on average per population. The mean effective number of alleles (Ae) and genetic diversity (H_e) were 2.4 and 0.418, respectively. The geographical distributions of corrected allele richness (Ar_5) and genetic diversity for each population are shown in Fig. 4. Our individual identification analysis revealed that the data for 10 nSSR loci can differentiate two

siblings from the same population, and the likelihood of wrong ascertainment is less than 0.01 (Fig. S2). Thus, our individual identification analysis indicates that we took no repeated samples from the same clone. AMOVA indicated significant genetic differentiation (F_{ST} = 0.15583, P < 0.01) across the sampling range (Table 1). The isolation-by-distance analysis detected statistically significant but weak correlation between pair-wise geographical distance and genetic differentiation (R^2 = 0.135, P = 0.010; Fig. S3) compared with reported correlations for other species studied in subtropical China (e.g. R^2 = 0.384, Li *et al*., 2012; R^2 = 0.6183, Sun *et al*., 2014a; R^2 = 0.4305, Sun *et al*., 2014b; R^2 = 0.3270, Gong *et al*., 2016).

The BARRIER analyses detected well-supported genetic and/or biogeographic barriers, separating populations into five groups (Figs. 5D and S4). AMOVA indicated that among-group, among-population and within-population variation accounted for 6.14%, 11.14% and 82.72% of the total genetic variation, respectively (Table S5). MIGRATE analyses of gene flow between groups indicated that little gene flow occurred between Group III and those to the north (IV and V; Fig. S5 and Table 3, S6). Indications of high levels of gene flow into the single-population groups (I and II) were obtained, but these may have been artifacts associated with very uneven sizes of the groups of populations (Table 3 and S6).

Population structure based on nuclear microsatellite markers

In the Bayesian clustering analysis using STRUCTURE, results of the delta *K* method suggest that the genetic structure of *P. adenopoda* can best be explained by two clusters (Fig. S6b). However, four (*K* = 4) or six (*K* = 6) clusters could also reflect the true number for three reasons. First, of the five *K* (2-6) values with narrow standard deviations of log probability, delta *K* values for *K* = 4 and *K* = 6 also show peaks, albeit lower than the peak for *K* = 2. Second, the log probability is highest for *K* = 6, although the log probability of *K* = 4 seemed to have reached a plateau. Third, the L'(*K*) value is highest for *K* = 2, slightly lower for *K* = 6 and markedly lower for *K* = 4. Hence, *K* = 2, 4, 6 may all be potential true cluster number (Fig. S6).

When two clusters were assumed (*K* = 2), every population contained a mixture of both clusters, but cluster 2 dominated populations of the Y-G Plateau (BARRIER Group III) and north-central part of the range, while cluster 1 was the major cluster in all northeastern and northwestern populations except P34 (Fig. 5A). When four or six clusters (*K* = 4 or *K* = 6) were assumed, Group III was again dominated by a single cluster (cluster 1 in both cases), whereas cluster 4 dominated most populations of north-central parts of the range (Fig. 5B, 5C). Populations in the northeast and northwest contained mixtures of various clusters other than cluster 1 under *K* = 4 or *K* = 6. Local subgroups showing similar proportions of clusters were discernible, especially under *K* = 6: P29, P30 and P31 in the far northwest had high proportions of cluster 5, whereas the nearby populations P26, P27 and P28 had higher proportions of clusters 3 and 4. The most northerly populations, P32 and P33, were dominated by

cluster 3, whereas the most easterly ones (P38, P39, P36 and P37) all had >50% cluster 2. However, these are largely differences in proportions of clusters (rather than presence or absence) among populations, and are consistent with gene flow among all these populations and groups thereof, except that there seemed to be little gene flow in or out of south-central Group III (Fig. 5). To facilitate discussion, the subsequent text concerns results for $K = 6$ unless otherwise stated, as they provide more details for visually inspecting differences and connections among populations.

Meanwhile, we detected no clear signs of sharing of genetic components between the 39 populations of *P. adenopoda* and 76 populations of the *P. Davidiana-rotundifolia* complex in our Bayesian clustering analysis when $K = 2$ (Fig. S7).

Inferring demographic history using approximate Bayesian computation (ABC)

The DIYABC analyses provide the strongest support for Sc6 (scenario 6, Fig. S8), i.e. an ancient range expansion, followed by a bottleneck event and a recent range expansion (Fig. 2). Both estimates of posterior probabilities (obtained from the direct approach and logistic regression) for this scenario (95.00%, 89.64%) are far higher than those for Sc1 (0.00%, 0.00%), Sc2 (0.00%, 0.00%), Sc3 (0.00%, 0.00%), Sc4 (5.00%, 10.35%) and Sc5 (0.00%, 0.00%). Moreover, Sc6 has the lowest average type I error rate (6.84%, Table S7). The marginal posterior probability densities for the demographic parameters of Sc6 (Table S8) indicate that *P. adenopoda* populations

initially had an effective population size (*N*e) of 20,700 (95% confidence interval: 11,300-29,100) individuals (N5; Fig. 6 and Table S8), and an expansion occurred 139.2 (54.8-218.0) Ka (time t2), increasing *N*e to 318,000 (69,100-570,000) individuals (N4). Next, a strong bottleneck occurred around 25.88 Ka (t1 + db) and ended at time t1, 20.24 (8.96-33.84) Ka, during which *N*e declined to 815 (171-1930) individuals (N2). *N*e subsequently rose to the current 244,000 (38,700-469,000) individuals (N1; Fig. 6 and Table S8). Despite the wide 95% confidence intervals, the initial expansion clearly seemed to occur around the beginning of the LIG, whereas the bottleneck respectively began and ended around the beginning and end of the LGM.

Species distribution modeling

The MAXENT models based on our field survey dataset (39 presence sites) generated robust predictions of *P. adenopoda*'s past, present, and future potential distributions. Average areas under the ROC curve (AUC) from 20 replicate MAXENT runs for the present-day, LGM, LIG, MH, 2050 and 2070 models of *P. adenopoda* distributions were 0.982 ± 0.005, 0.983 ± 0.006, 0.976 ± 0.012, 0.980 ± 0.004, 0.979 ± 0.005 and 0.980 ± 0.005, respectively, indicating that our models differed greatly from random expectations.

The predicted current distributions of *P. adenopoda* (Fig. 7A) strongly overlapped with the population sampling range, indicating that our genetic survey covered the

core of the species' potential range in central and southwestern China. The predicted present-day and MH distributions are nearly identical (Fig. 7). However, the LGM model showed a clear southward and westward range contraction relative to the present day, with relatively low habitat suitability (< 0.50) for northern parts of the current range and eastern China, but moderate habitat suitability in central and southwestern China (0.50-0.75; Fig. 7). The LIG model also showed a clear southward range contraction in central China relative to the present day, but expansion in eastern China (Fig. 7).

The potential distribution of *P. adenopoda* was modelled under two climate change scenarios, for the years 2050 (model rcp45) and 2070 (model rcp85) (Fig. S9). Both scenarios showed similar distributions, based on areas with putatively moderate to high habitat suitability (> 0.5), to the present-day model (Fig. S9).

Encouragingly, the MAXENT models based on the field survey dataset (39 presence sites) and those based on herbarium samples in addition to this dataset (81 presence sites) generated largely similar potential distributions (Fig. S10). Finally, MAXENT distributions based on our field survey dataset and habitat suitability > 0.75 captured the species' major distribution area best, so we consider them in further discussion.

DISCUSSION

Genetic diversity and genetic structure

Our genetic survey of *P. adenopoda* across its range revealed several distinctive features relative to reported patterns for other tree species distributed in subtropical China. The number of cpDNA haplotypes (13, based on four cpDNA fragments) and the total cpDNA haplotype diversity (H_T = 0.502) were both low compared to those of many other tree species examined in the region, e.g. *Loropetalum chinense* (three cpDNA regions, H_T = 0.767; Gong *et al.*, 2016, q.v. for a brief summary), *Quercus glauca* (three cpDNA regions, H_T = 0.908; Xu *et al.*, 2015), and *Taxus wallichiana* (one cpDNA region, H_T = 0.884; Gao *et al.*, 2007). We also detected moderate nuclear microsatellite gene diversity (ten nSSRs, H_e = 0.418, Table S1), lower than the reported nuclear gene diversity of *Cercidiphyllum japonicum* (seven nSSRs, H_e = 0.838, Qi *et al.*, 2012), but higher than the nuclear gene diversity of *Loropetalum chinense* (AFLPs, H_e = 0.263; Gong *et al*., 2016) and *Emmenopterys henryi* (AFLPs, H_e = 0.217, mean value = 0.117; Zhang *et al*., 2016).

Genetic variation was unevenly distributed across *P. adenopoda*'s range. Three to five cpDNA haplotypes were detected in eight of the 39 populations, mostly located in the southern part of the sampling range (Fig. 3). Conversely, average corrected allelic richness (Ar_5) and expected heterozygosity (H_e) at nSSR loci were higher in some peripheral populations and northern populations (Table S1, Fig. 4). Havrdová *et al.* (2015) detected a similar pattern in European *Alnus glutinosa* populations.

AMOVA analysis detected a moderate level of genetic variation between populations, in both the examined cpDNA sequences (F_{ST} = 0.209) and nSSRs (F_{ST} =

0.156; Table 1). However, higher levels have been found in populations of many other tree species in subtropical China, e.g. *Ginkgo biloba* (cpDNA- and AFLP-based F_{ST} = 0.345 and 0.288, respectively; Gong *et al.*, 2008), *Eurycorymbus cavaleriei* (cpDNA-based F_{ST} = 0.929; Wang *et al.*, 2009), and *Loropetalum chinense* (cpDNA and AFLP-based F_{ST} = 0.705 and 0.226, respectively; Gong *et al.*, 2016). This, plus the wide range of the commoner cpDNA haplotypes, and lack of phylogeographic structure ($N_{ST} < G_{ST}$, $P > 0.05$), indicates that levels of gene flow and migration have been higher among *P. adenopoda* populations than among populations of the other mentioned tree species in subtropical China. In contrast to the other species, *P. adenopoda* is wind-pollinated, deciduous and produces wind-dispersed seeds (Fang *et al.*, 1999), so detection of the higher rates of genetic exchange and migration in cpDNA as well as nSSRs suggests that seed dispersal was a major mediator of gene flow.

First, sampling of two or more ramets from the same clones may have biased the estimates, although we made efforts to minimize this risk (see Supplementary Data text for detailed discussion). Second, interspecific hybridization and subsequent introgression may lead to shared polymorphisms among species. However, we detected no clear signs of introgression between *P. adenopoda* and *P. davidiana* (Figs. S1 and S7), although we have observed hybrid individuals of these species (see Supplementary Data text for detailed discussion).

Potential glacial refugia indicated by cpDNA haplotype diversity, fossil records and SDMs

Locations of a species' possible glacial refugia can be inferred by integrating and reconciling three major lines of evidence: fossil records, SDMs and phylogeographic data (for a review, see Gavin *et al*., 2014). According to a paleo-biome reconstruction based on pollen records, but with sparse sampling sites (Harrison *et al*., 2001), temperate deciduous and mixed forests in central and southern China have moved northward more than 1,000 km since the LGM. Warm temperate evergreen forests have also apparently migrated northward (Harrison *et al*., 2001; Qiu *et al*., 2011). As *P. adenopoda* occurs in warm temperate evergreen forests and the southern parts of temperate deciduous forests (Fang *et al.*, 1999), the paleo-biome reconstruction indicates that glacial refugia of this poplar may have been located in southern parts of its current range (Harrison *et al*., 2001). Similarly, our SDMs showed a southward range contraction of *P. adenopoda* in the LGM (Fig. 7C) relative to the current distribution (Fig. 7A). Only the central and southern parts of its current distribution range were predicted to provide moderately to highly suitable habitats during the LGM (Fig. 7C). Thus, the SDMs suggest that these areas provided potential refugia for the species during glacial periods.

In phylogeographic studies, glacial refugia are usually detected by the presence of high diversity of haplotypes and major lineages within a species' populations (Comes and Kadereit, 1998; Hewitt, 2000). We detected three or more haplotypes in only

eight of the examined *P. adenopoda* populations, seven of which (P1, P2, P10, P12, P13, P16, and P17) were in the southern part of its range, the eastern Y-G Plateau (Fig. 3). Hence, this area appears to have been the most important for preservation of *P. adenopoda*'s haplotype diversity during glacial periods, i.e. the location of its main refugia. This plateau also putatively provided glacial refugia for *Fagus longipetiolata*, *Pinus kwangtungensis*, and *Taxus wallichiana* (Gao *et al.*, 2007; Liu, 2008; Tian *et al.*, 2008). Private haplotypes, found only in one population, may also be indicative of minor refugia (microrefugia or cryptic refugia) during glacial periods (e.g. Opgenoorth *et al.*, 2010). P18, P20, P22 and P24 in the central part of *P. adenopoda*'s range (Fig. 3) contain four private haplotypes and lie outside the putative major glacial refugia (P22 also contains two other haplotypes), and thus may occupy sites of such minor refugia. These populations are located in the northern Wuling Mountains and eastern Dabashan Mountains. The latter are putative glacial refugia for a group of tree species that occur or mainly occur in subtropical China, e.g. *Quercus glauca* (Xu *et al.*, 2015), *Eurycorymbus cavaleriei* (Wang *et al.*, 2009), and *Taxus wallichiana* (Gao *et al.*, 2007).

Strong bottleneck during the late Quaternary

In agreement with the SDMs (Fig. 7) and paleo-biome reconstruction (Harrison *et al.*, 2001), discussed above, indicating or predicting that *P. adenopoda*'s range contracted southward during the LGM, our coalescent-based demographic analyses

(DIYABC) of nSSR data indicated that the species experienced a very strong bottleneck during the LGM (Figs. 2 and 6). The likeliest scenario (Sc6) indicated that the species' effective population size shrank ca. 390-fold to just 815 individuals (N2) at around 25.88 Ka (t1 + db), from 318,000 before the bottleneck event (N4), and the population size increased again at 20.24 Ka (t1). Despite the uncertainty of estimates derived from coalescent-based demographic analyses in terms of 95% confidence intervals (Table S8) and generation times (Table S9), the apparent timing of bottleneck event correspond well with the beginning and end of the LGM, around ca. 26.5 Ka and ca. 20-19 Ka, respectively (Clark *et al*., 2009). The estimated effective size of the ancestral population size of *P. adenopoda* during the bottleneck event (815 individuals) is much smaller than reported estimates for ancestral populations of *P. tremula* (~118,000 individuals; Ingvarsson, 2008) and *P. balsamifera* (~20,000 individuals; Keller *et al*., 2010). However, estimates in the two cited studies were based on nuclear DNA sequence data, rather than nSSR sequences, so the apparent differences in *Ne* during the bottleneck should be interpreted cautiously.

During a bottleneck event, the estimated effective population size is smaller and genetic drift stronger than normal, which may result in increased frequencies of common alleles, decreased frequencies of rare alleles, and thus overall reductions in allele richness and genetic diversity (Freeland *et al*., 2011).

Accordingly, in agreement with our DIYABC analyses, we found relatively low allele richness at the examined nSSR loci: both the mean number of alleles per locus

(A = 3.5) and mean effective number of alleles per locus (Ae = 2.4) we obtained are lower than corresponding estimates for trembling aspen in North America (*P. tremuloides*, A = 6.4, Ae = 3.5, 12 nSSR loci; Latutrie *et al*., 2016; Ae = 5.99, eight nSSR loci; Callahan *et al*., 2013), white aspen in Europe (more specifically, Sardinia and Ticino: Ae = 3.61 and 4.17, respectively, five SSR loci; Brundu *et al*., 2008), and two desert poplars in northwestern China (*P. euphratica* and *P. pruinose*: Ae = 3.68 and 2.86, respectively, eight SSR loci; Wang *et al*., 2011).

The cpDNA genetic diversity of *P. adenopoda* also indicates that the species may have experienced a population bottleneck. As discussed above, we detected 13 cpDNA haplotypes in all 371 trees from 39 populations, and the total haplotype diversity (H_T) was 0.502. Five of the haplotypes (H1, H2, H3, H6, H12) were detected in ca. 95% of the sampled trees (352 out of 371); the other eight were detected in the other 29 sampled trees, three of them in singletons (H4, H10, H11; Table S1). This is lower than published diversity estimates for other poplars with similar life histories and mating systems to *P. adenopoda*. For example, a PCR-RFLP analysis of six cpDNA fragments in 414 individuals from 22 populations of *P. alba* and *P. tremula* in Central Europe detected 59 haplotypes, and provided total haplotype diversity estimates for the species of 0.898 and 0.905, respectively (Fussi *et al*., 2010). Another genetic survey detected seven haplotypes (based on unique event polymorphisms in two cpSSRs and one cpDNA fragment) solely in the *P. alba* population of the Iberian Peninsula (Macaya-Sanz *et al*., 2012). However, in North America, *Populus*

balsamifera (occupying areas that were nearly completely covered by glaciers during the LGM) reportedly harbors a similar number of cpDNA haplotypes to *P. adenopoda*, although it has a much larger range (Breen *et al*., 2012).

Evidence for repeated range contractions and expansions

Both the mismatch distribution analysis of the cpDNA haplotypes (Table 2) and star-like haplotype network obtained (Fig. 3) clearly indicate that demographic expansion events occurred in the evolutionary history of *P. adenopoda* (Avise, 2004; Peng *et al*., 2015). Our coalescent-based analyses of nSSR data also indicate that *P. adenopoda* probably experienced at least two rounds of range expansion and one bottleneck event (Fig. 6). The first range expansion event, involving a more than ten-fold increase in population size around 139.2 Ka, roughly corresponded to the beginning of the LIG, 131 to 114 Ka. A second ca. 300-fold population expansion from N2 = 815 individuals during the bottleneck to N1 = 244,000 individuals apparently occurred as *P. adenopoda* geographically spread to its current distribution. This expansion seems to have started at 20.24 Ka, closely matching the end of the LGM around 20-19 Ka (Clark *et al*., 2009). However, this timing should be treated cautiously, as correlations between demographic and climatic events (e.g. the LIG and LGM) are weak if the generation time of *P. adenopoda* is assumed to be shorter or longer than the 40 years assumed for these calculations (e.g. 20 years or 60 years, see Table S9).

As a strong recent bottleneck and subsequent range expansion may have partly or wholly obscured signatures of earlier demographic events and nSSRs have rapid mutation rates, we may not be able to detect bottleneck and range expansion events in earlier glacial and interglacial periods by coalescent-based analyses of nSSR data alone. However, further integration of the SDMs and inferred demographic dynamics during the LGM and LIG may be helpful to address this issue. DIYABC analysis indicated that *P. adenopoda* experienced a strong bottleneck during the LGM, and SDMs predicted that its range during the LGM (Fig. 7C) was narrower than its ranges both now (Fig. 7A) and during the LIG (Fig. 7D). The SDMs also indicate that the species' populations expanded both during the LIG and post-glacially, and its predicted range during the LIG (Fig. 7D) is similar to its range now (Fig. 7A) and during the Middle Holocene (Fig. 7B). It should be noted that although the species was not present in northern parts of its current range, a nearly continuous distribution belt in southern China may have formed during the LIG (Fig. 7D). Notably, our SDMs did not predict that anticipated climate change in the near future will substantially alter the species' distribution (Fig. S9). However, *P. adenopoda* likely underwent repeated demographic contractions and expansions during previous glacial and interglacial cycles (Petit *et al*., 1999; Lisiecki and Raymo, 2005).

Southern major glacial refugia made very limited contributions to postglacial recolonization in northern and eastern parts of the range

Both SDMs and the phylogeographic data indicate that the Y-G Plateau, in the southern part of *P. adenopoda*'s current range provided its main glacial refugia, as discussed above. However, our nSSR data appear to exclude most of the Y-G Plateau as a source of material for postglacial recolonization of northern and eastern parts of the range. The populations from the Y-G Plateau contain principally STRUCTURE cluster 1 (Fig. 5), which is very rare elsewhere, indicating very limited gene flow out of the Y-G Plateau. To confirm this pattern, we estimated gene flow between five groups of populations (separated by genetic discontinuities identified by BARRIER analyses, as shown in Figs. 5D and S4), using MIGRATE. According to both models, there was very limited historical gene flow between group III (mainly located on the Y-G Plateau) and either group IV (mainly located in the mountainous regions east and north of the Sichuan Basin) or group V (located in eastern parts of the range) (Table 3, Fig. S5). Hence, major glacial refugia in the Y-G Plateau may have made very limited contributions to the postglacial recolonization of northern and eastern parts of the range.

However, there may have been minor glacial refugia (microrefugia and cryptic refugia) to the north of the major southern glacial refugia during the LGM, and they may have contributed significantly to postglacial recolonization by the northern group IV and group V. Several lines of evidence support this hypothesis. First, SDMs predicted the presence of moderately suitable habitats in the areas occupied by southern elements of group IV, and a few populations of group V, during the LGM.

Second, the STRUCTURE analysis indicates that populations in the north form discernible subgroups, with similar proportions of clusters, e.g. P29-31, P27-28, P18-23+P34, P32+33, and P36+37+38+39 (Fig. 5). Finally, many populations in groups IV and V have higher corrected allele richness (Ar_5) and expected heterozygosity (H_e) than populations in other groups (Table S1, Fig. 4), indicating admixture of different lineages that probably survived in different minor glacial refugia. MIGRATE analysis based on nSSR data suggests that gene flow from group IV to group V was stronger than flows in the opposite direction (Table 3, Fig. S5), also suggesting lineage admixture. Moreover, without lineage admixture, founder effects usually lead to lower allele richness and expected heterozygosity during range expansion (Freeland *et al*., 2011). A recent study on *Alnus incana* in Europe found evidence of a similar pattern, of postglacial recolonization of northern parts of the range from minor glacial refugia in central parts, rather than from major glacial refugia in the southern part (Mandák *et al*., 2016).

Two caveats regarding our genetic survey should be noted. First, recolonization of northern and eastern parts of *P. adenopoda*'s range could have started from undetected glacial refugia. For example, group V may have come from the Wuyi Mountains in southeastern China, which provide highly suitable habitats according to our present-day SDMs (Fig. 7A), and there are herbarium records of populations at various sites in the region. However, we have found no populations there, despite visiting these sites several times. This area has a dense human population, some of

these *P. adenopoda* populations were eliminated by changes in land use (e.g. the population at Longjing, Zhejiang, was destroyed by a farmer cutting down the aspens to make space for cultivating tea), and others probably met several fates. Nevertheless, future genetic surveys of the species should include samples from peripheral parts of the range, as well as the known sampling sites we used in this study. The other caveat is that, although our sampling covered most of the range of *P. adenopoda*, our inferences were based on a limited number of DNA markers (four cpDNA fragments and 10 nSSR loci). We believe we have captured the predominant pattern of the Quaternary population evolutionary history of *P. adenopoda*, but future population genomic surveys will probably provide a clearer picture.

In summary, our integrative approach revealed that *Populus adenopoda* has had a complex evolutionary history in China. SDMs clearly suggested a southward contraction during the LGM, and low haplotype diversity in the north of the present range fits a model of postglacial expansion. All our data indicate that the Y-G Plateau provided major glacial refugia, but populations in minor glacial refugia in central parts of the species' range may have contributed substantially more to its postglacial range expansion. Our DIYABC simulations support the hypothesis that the species has experienced at least one expansion, a bottleneck and second expansion roughly coinciding with initiation of the LIG, start of the LGM and end of the LGM, respectively. The species' populations may have experienced more dynamic distributional changes than most of the other tree species in subtropical China. Our

study, and other recent studies, emphasize the importance of combining multiple lines of evidence, including (*inter alia*) distributions of uni-and bi-parentally inherited markers, coalescent-based analyses, SDMs, fossil records and related analyses, when inferring Quaternary population evolutionary histories of species in topographically complex regions.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. Fig. S1: Phylogenetic tree based on 13 cpDNA haplotypes (H1-H13) of *P. adenopoda* and 21 cpDNA haplotypes (H1-H21) of the *P. davidiana-rotundifolia* complex. Fig. S2: Comparison of the cumulative probability of identity for unrelated individuals (P (ID)), and siblings (P (ID) sib), P (ID) < 0.01.

Fig. S3: Isolation-by-distance plots showing the correlation between genetic distance [y-axis: $F_{ST}/(1–F_{ST})$], based on nSSR data, and geographical distance [x-axis: natural geographical distance] between the 39 sampled populations. Fig. S4: Major genetic discontinuities in nSSR genotypes of populations in different sampling areas inferred by Monmonier's maximum-difference algorithm implemented in BARRIER. Fig. S5: Illustrative representation of the effective population size (N_e) and effective migration rates (N_em) of the five BARRIER-inferred population groups under two hypotheses: (a) variable theta and (b) same theta. Fig. S6: Results of Bayesian clustering analysis of nSSR data to determine the optimum number of clusters (*K*) for the 39 sampled

populations. Fig. S7: Results of Bayesian clustering analysis of nSSR data to determine the optimum number of clusters (*K*) for both the 39 sampled populations of *P. adenopoda* and 76 sampled populations of the *P. davidiana-rotundifolia* complex. Fig. S8: Plots showing fitness of competing scenarios (Sc1-Sc6), based on direct estimates and logistic regression, simulated in DIYABC. Fig. S9: Potential distributions of *P. adenopoda* under future global warming scenarios derived, using the field survey dataset (39 presence sites), from species distribution models implemented in MAXENT. Fig. S10: Distribution dynamics of *P. adenopoda* from the late Quaternary to the present derived, using the field survey and herbarium sample dataset (81 presence sites), from species distribution models implemented in MAXENT. Table S1: Details of sample locations, sample sizes and descriptive statistics of genetic variability for the 39 populations of *Populus adenopoda*. Table S2: Primers used for amplifying and sequencing: (a) cpDNA regions and (b) nSSR loci. Table S3: Prior settings for all parameters used when simulating DIYABC scenarios. Table S4: Sequence variations among the 13 haplotypes (H1–H13) detected in the four examined cpDNA regions. Table S5: Results of analysis of molecular variance (AMOVA) of groups detected by BARRIER and populations of *Populus adenopoda* based on nSSR data. Table S6: Effective population sizes (N_e) of *Populus adenopoda*, and effective migration rates (N_em) between the five groups of the species' populations. Table S7: Estimates of type I and type II error probabilities for the six scenarios in DIYABC. Table S8: Estimates of posterior distributions of parameters

obtained from the Approximate Bayesian Computation (ABC) for the best scenarios for the demographic history of *Populus adenopoda*, with time parameters (t and db, unit years) scaled by an assumed generation time of 40 years. Table S9: Estimates of posterior distributions of parameters obtained from the Approximate Bayesian Computation (ABC) for the best scenarios for the demographic history of *Populus adenopoda*, with time parameters (t and db, unit years) scaled by assumed generation times of 20 and 60 years.

ACKNOWLEDGEMENTS

This study was funded by the National Natural Science Foundation of China (grant nos. 31590821, 31622015, 41571054), the National Basic Research Program of China (grant 2014CB954100), Sichuan Provincial Department of Science and Technology (grant 2015JQ0018) and Sichuan University. We thank Dr. Matthew Olson, Dr. Markus Rhusam, Dr. Jianquan Liu, and four anonymous reviewers for their constructive suggestions.

LITERATURE CITED

Amos W, Hoffman J, Frodsham A, Zhang L, Best S, Hill A. 2007. Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes* **7**: 10–14.

Avise JC. 2004. *Molecular makers, natural history, and evolution.* Sunderland, MA: Sinauer.

Avise JC. 2009. Phylogeography: retrospect and prospect. *Journal of Biogeography* **36**: 3–15.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37–48.

Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution and Systematics* **41**: 379–406.

Beerli P, Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* **152**: 763–773.

Beerli P, Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* **185**: 313–326.

Bennett K, Tzedakis P, Willis K. 1991. Quaternary refugia of north European trees. *Journal of Biogeography* **18**: 103–115.

Breen AL, Murray, DF, Olson MS. 2012. Genetic consequences of glacial survival: the late Quaternary history of balsam poplar (*Populus balsamifera* L.) in North America. Journal of Biogeography **39:** 918–928.

Brundu G, Lupi R, Zapelli I, *et al*. 2008. The origin of clonal diversity and structure of *Populus alba* in Sardinia: evidence from nuclear and plastid microsatellite markers. *Annals of Botany* **102:** 997–1006.

Callahan CM, Rowe CA, Ryel RJ, Shaw JD, Madritch MD, Mock KE. 2013. Continental-scale assessment of genetic diversity and population structure in quaking aspen (*Populus tremuloides*). *Journal of biogeography* **40:** 1780–1791.

Carstens BC, Richards CL. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* **61**: 1439–1454.

Clark PU, Dyke AS, Shakun JD, C *et al*. 2009. The last glacial maximum. *Science* **325:** 710–714.

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

Cornuet JM, Pudlo P, Veyssier J, *et al*. 2014. DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30**: 1187–1189.

Dakin EE, Avise JC. 2004. Microsatellite null alleles in parentage analysis. *Heredity* **93**: 504–509.

Dieringer D, Schlötterer C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3**: 167–169.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.

Editorial Committee of Vegetation Map of China, 2007. Vegetation Map of the People's Republic of China (1:1 000 000). Geology Publishing House, Beijing, China

Elith J, Leathwick JR. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* **40**: 677.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

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

Fang Z, Zhao S, Skvortsov A. 1999. Salicaceae. *Flora of China* **4**: 139–274.

Freeland JR, Kirk H, Petersen S. 2011. Molecular markers in ecology. *Molecular Ecology* 35-75. John Wiley & Sons, West Sussex, UK.

Fussi B, Lexer C, Heinze B. 2010. Phylogeography of *Populus alba* (L.) and *Populus tremula* (L.) in Central Europe: secondary contact and hybridisation during recolonisation from disconnected refugia. *Tree Genetics & Genomes* **6:** 439–450.

Gao L, Möller, M, Zhang XM, *et al*. 2007. High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Molecular Ecology* **16**: 4684–4698.

Gavin DG, Fitzpatrick MC, Gugger PF, *et al*. 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist* **204**: 37–54.

Gong W, Chen C, Dobeš C, Fu CX, Koch MA. 2008. Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion. *Molecular Phylogenetics and Evolution* **48**: 1094–1105.

Gong W, Liu W, Gu L, Kaneko S, Koch MA, Zhang D. 2016. From glacial refugia to wide distribution range: demographic expansion of *Loropetalum chinense* (Hamamelidaceae) in Chinese subtropical evergreen broadleaved forest. *Organisms Diversity & Evolution* **16**: 23–38.

Harrison S, Yu G, Takahara H, Prentice I. 2001. Palaeovegetation (Communications arising): diversity of temperate plants in east Asia. *Nature* **413**: 129–130.

Havrdová A, Douda J, Krak K, *et al*. 2015. Higher genetic diversity in recolonized areas than in refugia of *Alnus glutinosa* triggered by continent-wide lineage admixture. *Molecular Ecology* **24**: 4759–4777.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **359:** 183–195.

Hickerson M, Carstens B, Cavender-Bares J, *et al*. 2010. Phylogeography's past, present, and future: 10 years after. *Molecular Phylogenetics and Evolution* **54**: 291–301.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Hijmans RJ, Guarino L, Mathur P. 2012. DIVA-GIS. Version 7.5. A geographic information system for the analysis of species distribution data. Manual available at http://www.diva-gis.org.

Ingvarsson PK. 2008. Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula* Genetics **180:** 329–340.

Jiang B, Tan Y. 2009. Study on the Community Structure and Interspecies Competition of *Populus adenopoda*. *Journal of Fujian Forestry Science and Technology* **1**: 018.

Jiang DC, Feng JJ, Dong M, Wu GL, Mao KS, Liu JQ. 2016. Genetic origin and composition of a natural hybrid poplar *Populus* × *jrtyschensis* from two distantly related species. *BMC Plant Biology* **16**: 89.

Keller SR, Olson MS, Silim S, Schroeder W, Tiffin P. 2010. Genomic diversity, population structure, and migration following rapid range expansion in the Balsam Poplar, *Populus balsamifera*. *Molecular Ecology* **19:** 1212–1226.

Knowles LL, Maddison WP. 2002. Statistical phylogeography. *Molecular Ecology* **11**: 2623–2635.

Kou YX, Cheng SM, Tian S, *et al*. 2016. The antiquity of *Cyclocarya paliurus* (Juglandaceae) provides new insights into the evolution of relict plants in subtropical China since the late Early Miocene. *Journal of biogeography* **43**: 351–360.

Latutrie M, Bergeron Y, Tremblay F. 2016. Fine-scale assessment of genetic diversity of trembling aspen in northwestern North America. *BMC evolutionary biology* 16: 231.

López-Pujol J, Zhang FM, Sun HQ, Ying TS, Ge S. 2011. Centres of plant endemism in China: places for survival or for speciation? *Journal of Biogeography* **38**: 1267–1280.

Levsen ND, Tiffin P, Olson MS. 2012. Pleistocene speciation in the genus *Populus* (Salicaceae). *Systematic Biology* **61**: 401–412.

Li Y, Yan HF, Ge XJ. 2012. Phylogeographic analysis and environmental niche modeling of widespread shrub *Rhododendron simsii* in China reveals multiple glacial refugia during the last glacial maximum. *Journal of Systematics and Evolution* **50:** 362–373.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Liu M. 2008. *Phylogeography of* Fagus longipetiolata*: insights from nuclear DNA microsatellites and chloroplast DNA variation*. Dissertation, East China Normal University, Shanghai, China.

Lisiecki LE, Raymo ME. 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanography* **20**.

Lu H, Yi S, Liu Z, *et al*. 2013. Variation of East Asian monsoon precipitation during the past 21 ky and potential CO_2 forcing. *Geology* **41**: 1023–1026.

Ma T, Wang J, Zhou G, *et al*. 2013. Genomic insights into salt adaptation in a desert poplar. *Nature Communications* **4**: 2797.

Macaya-Sanz D, Heuertz M, LÓPEZ-de-HEREDIA U, *et al*. 2012. The Atlantic-Mediterranean watershed, river basins and glacial history shape the genetic structure of Iberian poplars. *Molecular Ecology* **21**: 3593–3609.

Mandák, B, Havrdová A, Krak K, *et al*. 2016. Recent similarity in distribution ranges does not mean a similar postglacial history: a phylogeographical study of the boreal tree species *Alnus incana* based on microsatellite and chloroplast DNA variation. *New Phytologist* **210:** 1395–1407.

Manni F, Guerard E, Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology* **76**: 173–190.

Opgenoorth L, Vendramin GG, Mao, KS, *et al*. 2010. Tree endurance on the Tibetan Plateau marks the world's highest known tree line of the Last Glacial Maximum. *New Phytologist* **185**: 332–342.

Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288–295.

Peng YL, Tian B, Tian XM, Wang J, Hensen I, Liu JQ. 2015. Range expansion during the Pleistocene drove morphological radiation of the fir genus (*Abies*, Pinaceae) in the Qinghai-Tibet Plateau and Himalayas. *Botanical Journal of the Linnean Society* **179**: 444–453.

Petit JR, Jouzel J, Raynaud D, *et al*. 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399:** 429–436.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modelling* **190**: 231–259.

Pons O, Petit RJ. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**: 1237–1245.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Qi XS, Chen C, Comes HP, *et al*. 2012. Molecular data and ecological niche modelling reveal a highly dynamic evolutionary history of the East Asian Tertiary relict *Cercidiphyllum* (Cercidiphyllaceae). *New Phytologist* **196**: 617–630.

Qiu YX, Fu CX, Comes HP. 2011. 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. *Molecular Phylogenetics and Evolution* **59**: 225–244.

Schroeder H, Hoeltken A, Fladung M. 2012. Differentiation of *Populus* species using chloroplast single nucleotide polymorphism (SNP) markers–essential for comprehensible and reliable poplar breeding. *Plant Biology* **14**: 374–381.

Shafer A, Cullingham CI, Cote SD, Coltman DW. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* **19**: 4589–4621.

Shi MM, Michalski SG, Welk E, Chen XY, Durka W. 2014. Phylogeography of a widespread Asian subtropical tree: genetic east–west differentiation and climate envelope modelling suggest multiple glacial refugia. *Journal of Biogeography* **41**: 1710–1720.

Simmons MP, Ochoterena H. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic biology* **49**: 369–381.

Sun Y, Moore MJ, Yue L, *et al.* 2014a. Chloroplast phylogeography of the East Asian Arcto-Tertiary relict *Tetracentron sinense* (Trochodendraceae). *Journal of Biogeography* **41**: 1721–1732.

Sun Y, Hu H, Huang H, Vargas-Mendoza CF. 2014b. Chloroplast diversity and population differentiation of *Castanopsis fargesii* (Fagaceae): a dominant tree species in evergreen broad-leaved forest of subtropical China. *Tree genetics & genomes* **10:** 1531–1539.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725–2729.

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

Tian S, Lei SQ, Hu W, *et al.* 2015. Repeated range expansions and inter-/postglacial recolonization routes of *Sargentodoxa cuneata* (Oliv.) Rehd. et Wils.

(Lardizabalaceae) in subtropical China revealed by chloroplast phylogeography. *Molecular Phylogenetics and Evolution* **85**: 238–246.

Tuskan GA, Gunter LE, Yang ZK, Yin T, Sewell MM, DiFazio SP. 2004. Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Canadian Journal of Forest Research* **34**: 85–93.

Tzedakis P, Emerson B, Hewitt G. 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution* **28**: 696–704.

Valière N. 2002. GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes* **2**: 377–379.

Wang J, Gao P, Kang M, Lowe AJ, Huang H. 2009. Refugia within refugia: the case study of a canopy tree (*Eurycorymbus cavaleriei*) in subtropical China. *Journal of Biogeography* **36**: 2156–2164.

Wang J, Wu Y, Ren G, Guo Q, Liu J, Lascoux M. 2011. Genetic differentiation and delimitation between ecologically diverged *Populus euphratica* and *P. pruinosa*. *PLoS One* **6:** e26530.

Wu Z, Zhu Y, Jiang H. 1987. Yunnan vegetation. In. Science Press, Beijing.

Wu ZY. 1980. Vegetation of China. In. Science Press, Beijing.

Xu J, Deng M, Jiang XL, Westwood M, Song YG, Turkington R. 2015. Phylogeography of *Quercus glauca* (Fagaceae), a dominant tree of East Asian subtropical evergreen forests, based on three chloroplast DNA interspace sequences. *Tree Genetics & Genomes* **11:** 1–17.

Ying J. 2000. Species diversity and distribution pattern of seed plants in China. *Chinese Biodiversity* **9**: 393–398.

Young ND, Healy J. 2003. GapCoder automates the use of indel characters in phylogenetic analysis. *BMC Bioinformatics* **4**: 6.

Yu G, Chen X, Ni J, *et al*. 2000. Palaeovegetation of China: a pollen data-based synthesis for the mid-Holocene and last glacial maximum. *Journal of Biogeography* **27**: 635–664.

Zhang YH, Wang IJ, Comes HP, Peng H, Qiu YX. 2016. Contributions of historical and contemporary geographic and environmental factors to phylogeographic structure in a Tertiary relict species, *Emmenopterys henryi* (Rubiaceae). *Scientific reports* **6**: 24041.

Zheng HL, Fan LQ, Milne RI, Zhang L, Wang YL, Mao KS. 2017. Species delimitation and lineage separation history of a species complex of aspens in China. *Frontiers in plant science* **8**: 375.

Zheng Y, Yu G, Wang S, Xue B, Liu H, Zeng X. 2003. Simulations of LGM climate of East Asia by regional climate model. *Science in China Series D: Earth Sciences* **46**: 753–764.

Zhou YB, Chen WW, Kaneko YY, *et al*. 2015. Seasonal dietary shifts and food resource exploitation by the hog badger (*Arctonyx collaris*) in a Chinese subtropical forest. *European journal of wildlife research* **61**: 125–133.

Tables

Table 1. Analysis of molecular variance (AMOVA) for populations of *Populus adenopoda* based on cpDNA data and nSSR data.

Source of variation	df	SS	VC	V%	*F*-statistics
CpDNA					
Among populations	38	3.063	0.00607	20.93	F_{ST} = 0.20931*
Within populations	332	7.611	0.02293	79.07	
Total	370	10.674	0.02899		
Nuclear microsatellite					
Among populations	38	453.534	0.44985	15.58	F_{ST} = 0.15583*
Within populations	791	1927.667	2.43700	84.42	
Total	892	2381.201	2.88685		

Abbreviations: df, degrees of freedom; SS, sum of squares; VC, variance components; V%, percent variation; F_{ST}, differentiation among populations; *, $P < 0.01$, 1023 permutations.

Table 2. Summary results of demographic analyses based on cpDNA data.

			Mismatch distribution (demographic expansion)		Mismatch distribution (spatial expansion)		Neutrality test	
Species	θ_0	θ_1	SSD (*P*-value)	HRI (*P*-value)	SSD (*P*-value)	HRI (*P*-value)	Tajima's *D* (*P*-value)	Fu's *F*s (*P*-value)
P. adenopoda	0.00000	0.06367	0.00001 (0.155)	0.78586 (0.851)	0.00001 (0.186)	0.78586 (0.801)	-0.5768 (0.272)	-0.79739 (0.152)

Abbreviations: θ_0 and θ_1, pre- and post-expansion populations sizes; SSD, sum of squared deviations; HRI, Harpending's raggedness index; Fu's *F*s and Tajima's *D* are shown with *P* values as above.

Table 3. Effective population sizes (N_e) of *Populus adenopoda*, and effective migration rates (N_em) between each pair of the species' five population groups (indicated in Fig. 5D).

(a) Variable Theta

		$2N_e$m				
	N_e	$2N_e$m I→	$2N_e$m II→	$2N_e$m III→	$2N$em IV→	$2N_e$m V→
I	**67.50** (0-200)		**7.58** (0.000-32.911)	**8.24** (0.000-47.520)	**27.95** (0.000-115.499)	**18.42** (0.000-78.584)
II	**132.50** (0-255)	**1.09** (0.000-18.187)		**3.53** (0.000-38.280)	**25.29** (0.000-112.201)	**22.85** (0.000-86.856)
III	**267.50** (95-450)	**0.89** (0.000-17.413)	**1.36** (0.000-20.944)		**11.98** (0.000-95.699)	**5.16** (0.000-59.972)
IV	**907.50** (685-1125)	**0.50** (0.000-16.427)	**1.36** (0.000-20.944)	**2.75** (0.000-36.960)		**5.16** (0.000-59.972)
V	**502.50** (295-705)	**0.69** (0.000-16.427)	**1.36** (0.000-21.692)	**2.75**(0.000-38.280)	**22.63** (0.000-112.201)	

(b) Same Theta

		$2N_e$m				
	N_e	$2N_e$m I→	$2N_e$m II→	$2N_e$m III→	$2N_e$m IV→	$2N_e$m V→
I	**35.00** (0-126.7)		**1.10** (0.000-7.520)	**7.41** (0.000-22.489)	**22.89** (1.640-53.182)	**5.51** (0.000-25.800)
II	**78.30** (0-156.7)	**0.58** (0.000-6.418)		**7.41** (0.000-22.000)	**21.49** (0.000-53.182)	**27.83** (8.000-56.400)
III	**258.30** (146.7-366.7)	**0.21** (0.000-4.898)	**0.47** (0.000-6.267)		**5.99** (0.000-29.165)	**2.53** (0.000-18.600)
IV	**528.30** (410-643.3)	**0.16** (0.000-4.729)	**0.37** (0.000-5.849)	**1.55** (0.000-14.667)		**2.53** (0.000-20.400)
V	**345.00** (240-450)	**0.21** (0.000-5.067)	**1.10** (0.000-7.729)	**1.55** (0.000-14.178)	**13.03** (0.000-36.8855)	

N_e, effective population size.

μ, mutation rate (μ= 10^{-3} per gamete per year).

$2N_e$m I→, the effective number of migrants from group I to groups II, III, IV, V, respectively.

$2N_e$m II→, the effective number of migrants from group II to groups I, III, IV, V, respectively.

$2N_e$m III→, the effective number of migrants from group III to groups I, II, IV, V, respectively.

$2N_e$m IV→, the effective number of migrants from group IV to groups I, II, III, V, respectively.

$2N_e$m V→, the effective number of migrants from group V to groups I, II, III, IV, respectively.

The mode value of the posterior distribution of each parameter is listed, and values of the lower and upper 95% credibility intervals are shown in brackets.

Figure legends

Fig. 1. Map of the presumed geographic distribution of *P. adenopoda* and the sampling locations of 39 populations addressed in this study (black circles). The code and provenance of each population are listed in Table S1. Our sampling covered the main part of the species' potential range, see also Fig. 7.

Fig. 2. Schematic representation of the six demographic scenarios (including model parameters) for *P. adenopoda* tested by approximate Bayesian computation (ABC) using the nuclear microsatellite data: (Sc1) constant effective population size (N5), (Sc2) an old expansion (N5 to N1, t2), (Sc3) a recent expansion (N5 to N1, t1), (Sc4) an old bottleneck (N5 to N2, t2) followed by a recent expansion that led to current effective size (N2 to N1, t1), (Sc5) an old expansion (N5 to N3, t2) followed by a recent bottleneck (N3 to N2, t1), and (Sc6) an old expansion (N5 to N4, t2) followed by a bottleneck (N4 to N2, t1+db), and then a recent expansion (N2 to N1, t1). For settings of demographic parameters, see Table S3. Times and effective population sizes are not strictly to scale.

Fig. 3. Map of the sampling locations and geographic distribution of cpDNA haplotypes. The code and provenance of each population are listed in Table S1. The bottom central inset depicts a median joining phylogenetic network showing phylogenetic relationships among the cpDNA haplotypes; all branches indicate a

single mutational step except for that connecting H6 to the outgroup (22 mutations). Diameters of the circles are proportional to the numbers of samples per haplotype.

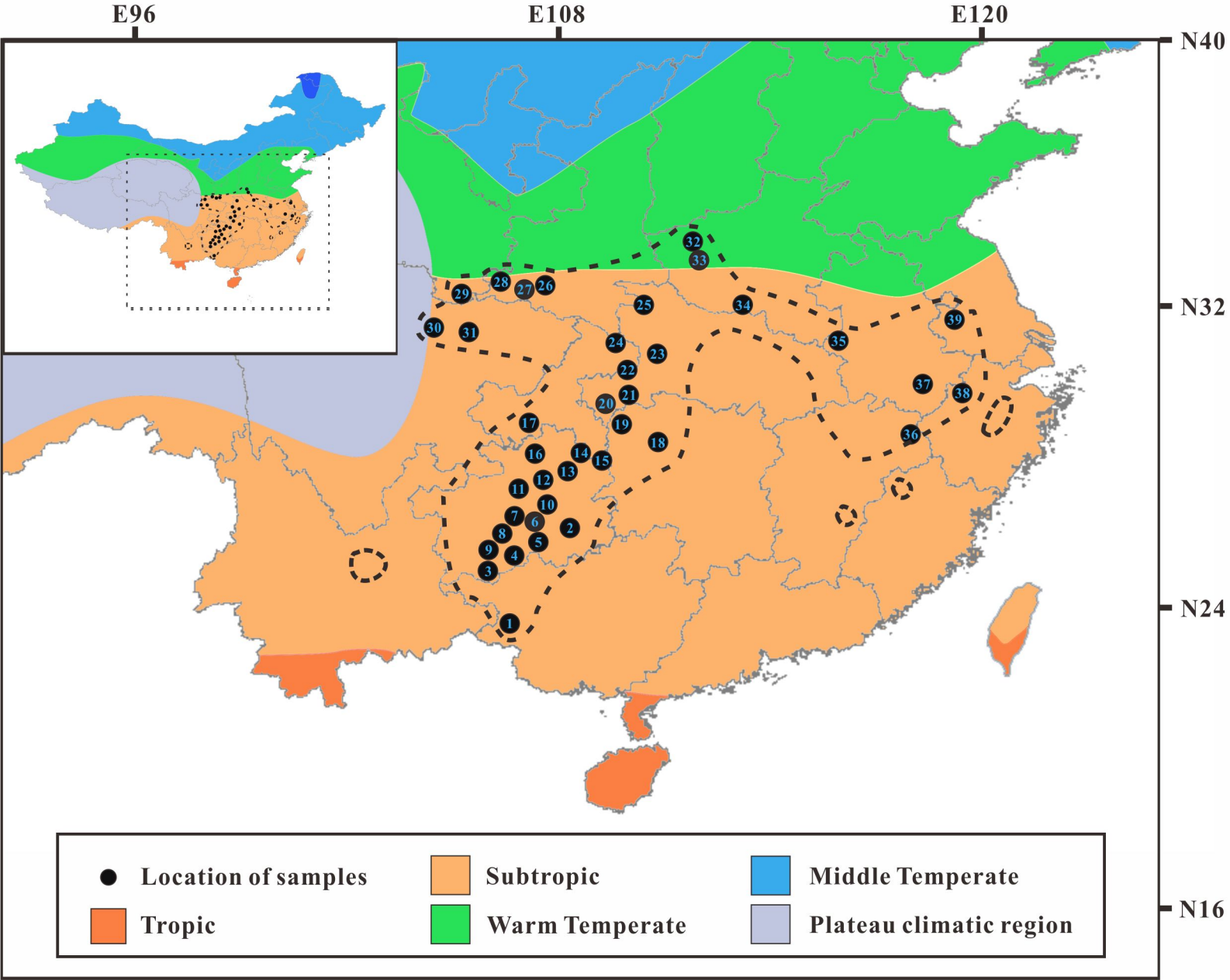
Fig. 4. Maps of (A) allelic richness (Ar_5) and (B) expected heterozygosity (H_e) for each sampled population. Geographical distributions of genetic diversity were visualized by mapping variation in allelic richness (Ar_5, based on five samples per population) and expected heterozygosity (H_e) across 39 sampled populations of *P. adenopoda* in DIVA-GIS v. 7.5 (Hijmans *et al.*, 2012).

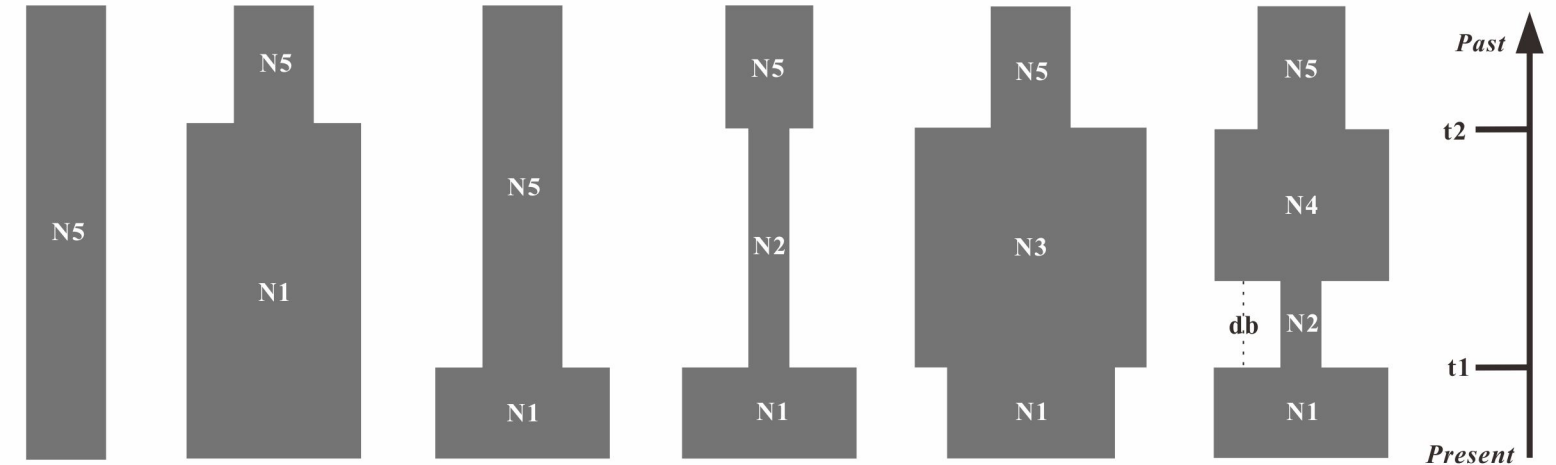
Fig. 5. Map of the geographic distribution of nuclear microsatellite clusters (inferred by STRUCTURE) when the assumed cluster numbers are (A) $K = 2$, (B) $K = 4$, and (C) $K = 6$, and (D) population groups (inferred by BARRIER) for the 39 sampled populations of *P. adenopoda*. Color coded regimes for (A), (B) and (C) are the same as in Fig. S5.

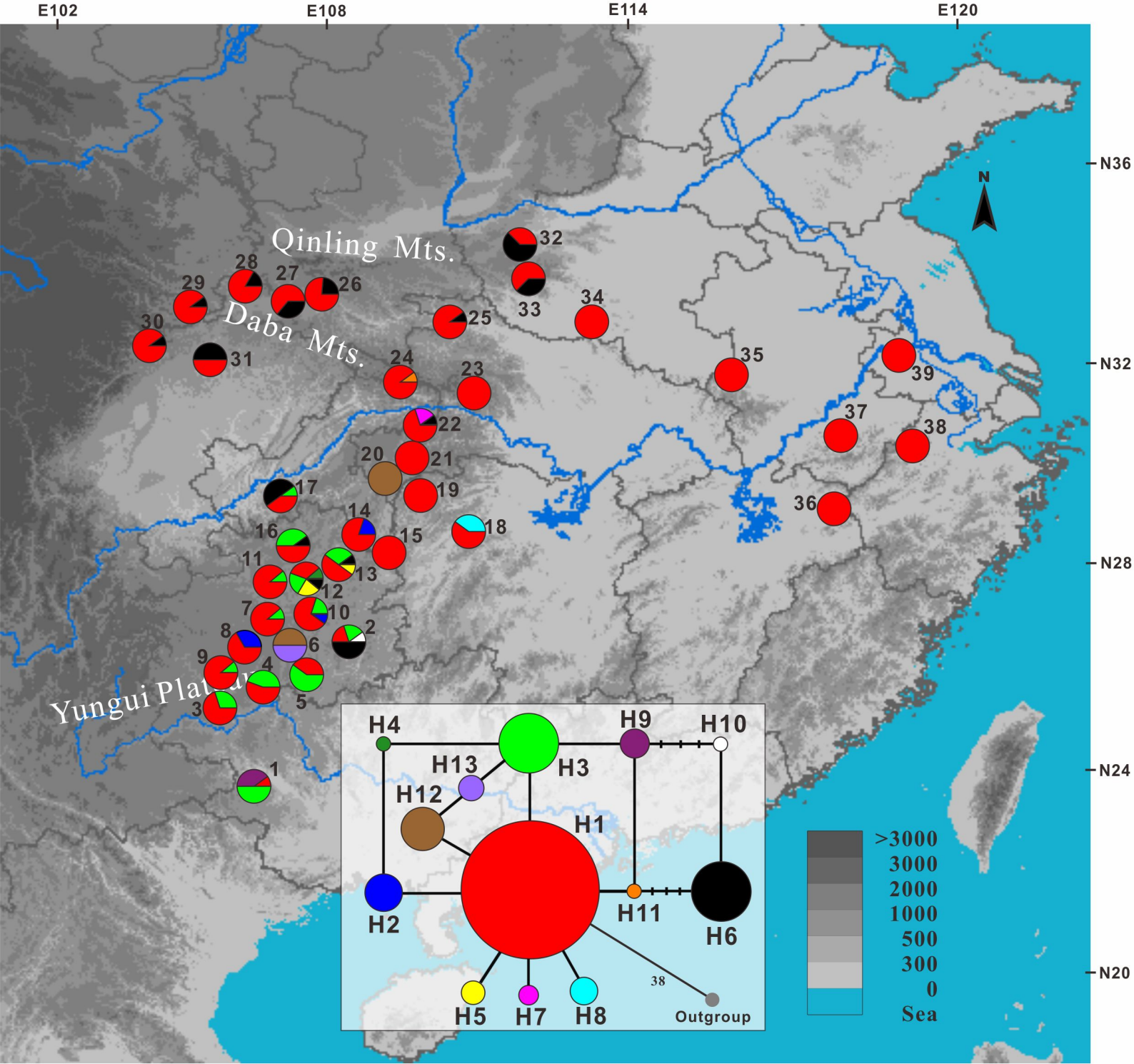
Fig. 6. Schematic representation for the estimates of effective population size (N_e) and the N_e-transition time of the best DIYABC scenario (Sc6). Black vertical bars and white horizontal bars represent 95% confidence interval for the estimates of N_e and N_e-transition time, respectively. Note that the y-axis of N_e value is in Log_{10} format.

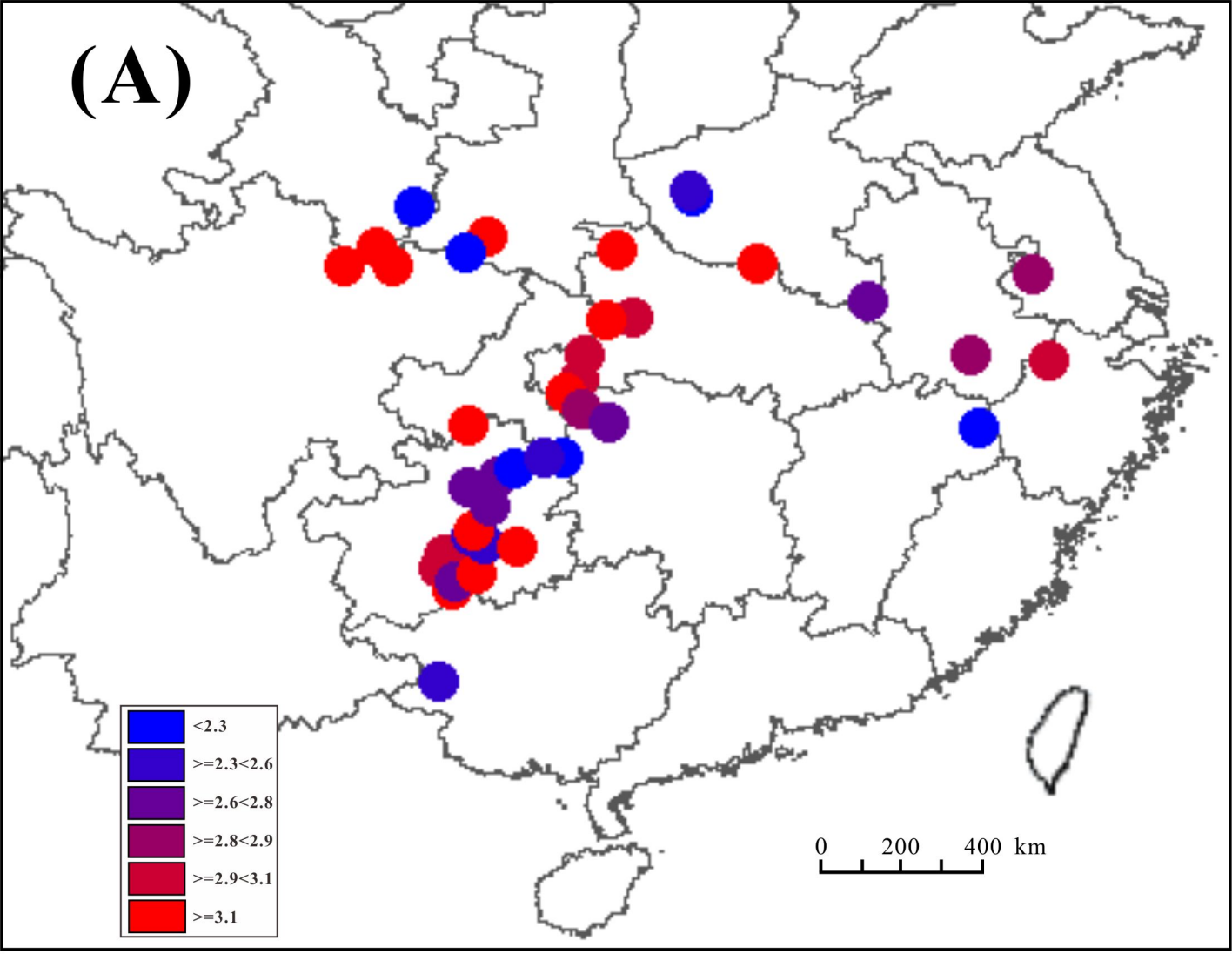
Fig. 7. Distribution dynamics of *P. adenopoda* from the late Quaternary to the present based on ecological niche modelling using MAXENT. Predicted distributions are shown for: (A) the present-day, (B) the MH, (C) the LGM and (D) the LIG models. For each of the four time periods, models were simulated for 20 replicates using 80%

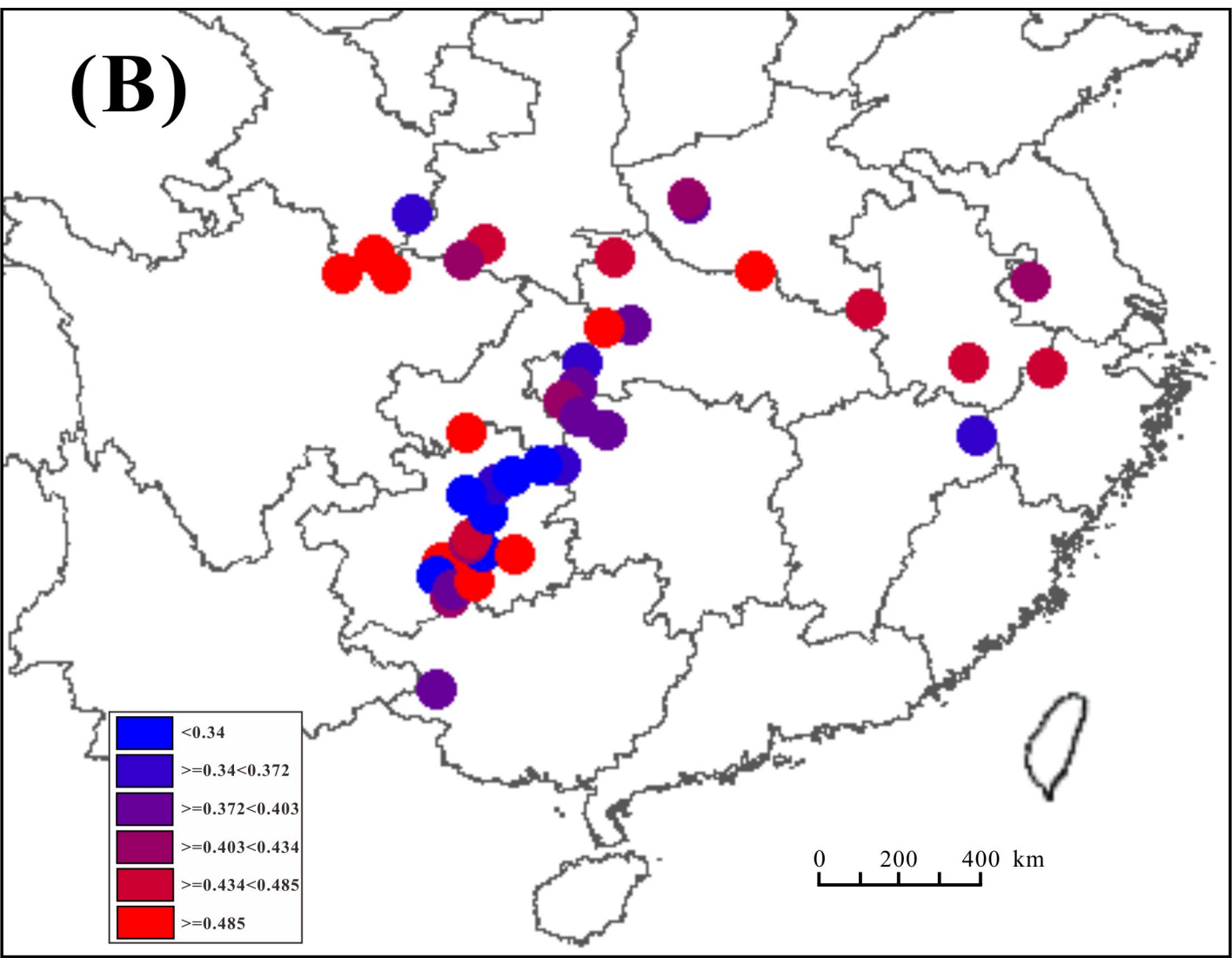
of the distribution coordinates for training and 20% for testing, and the average-over-all-replicates model was adopted.

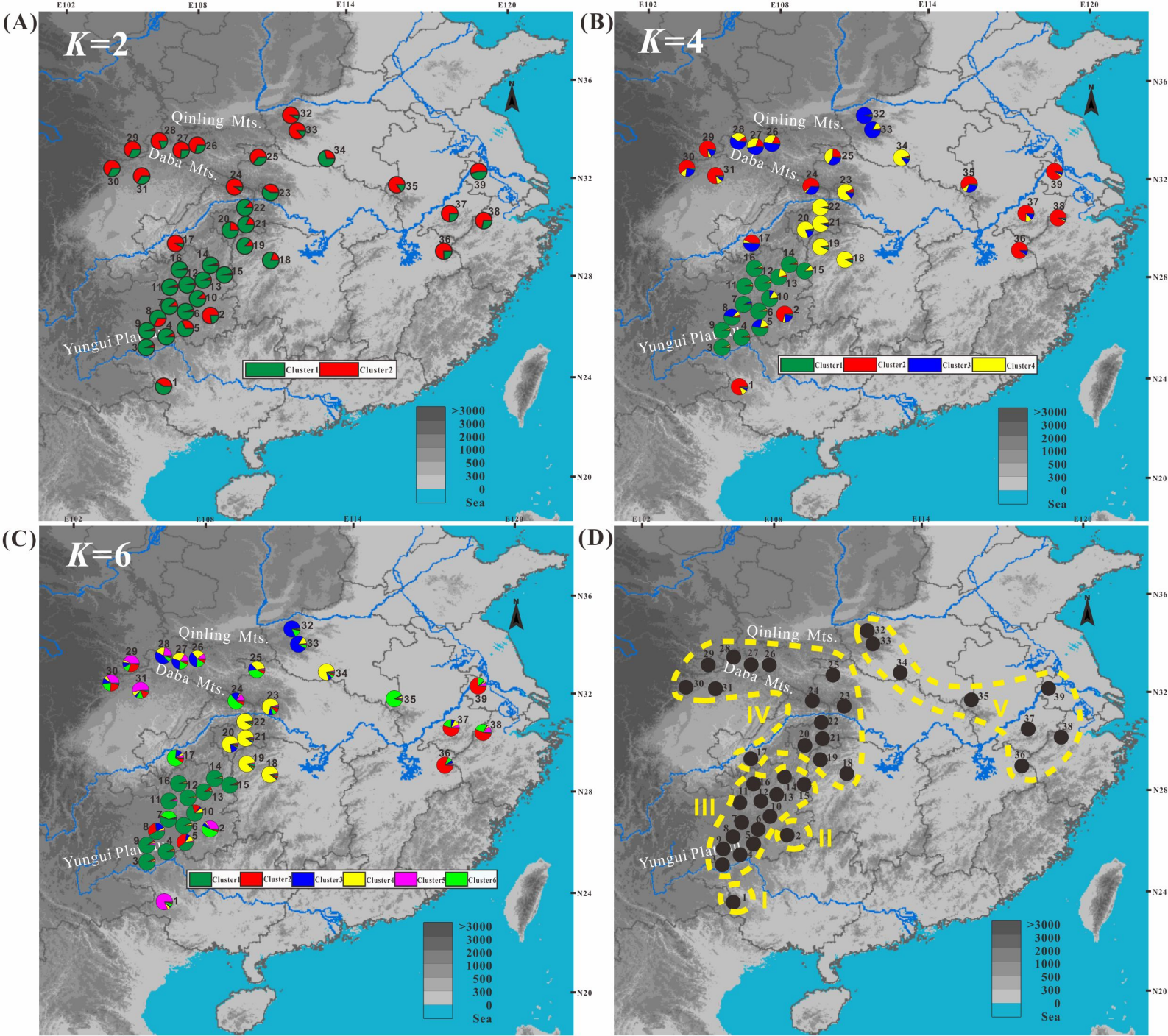

E96
E108
E120
N40
N32
N24
N16
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
Location of samples
Subtropic
Middle Temperate
Tropic
Warm Temperate
Plateau climatic region

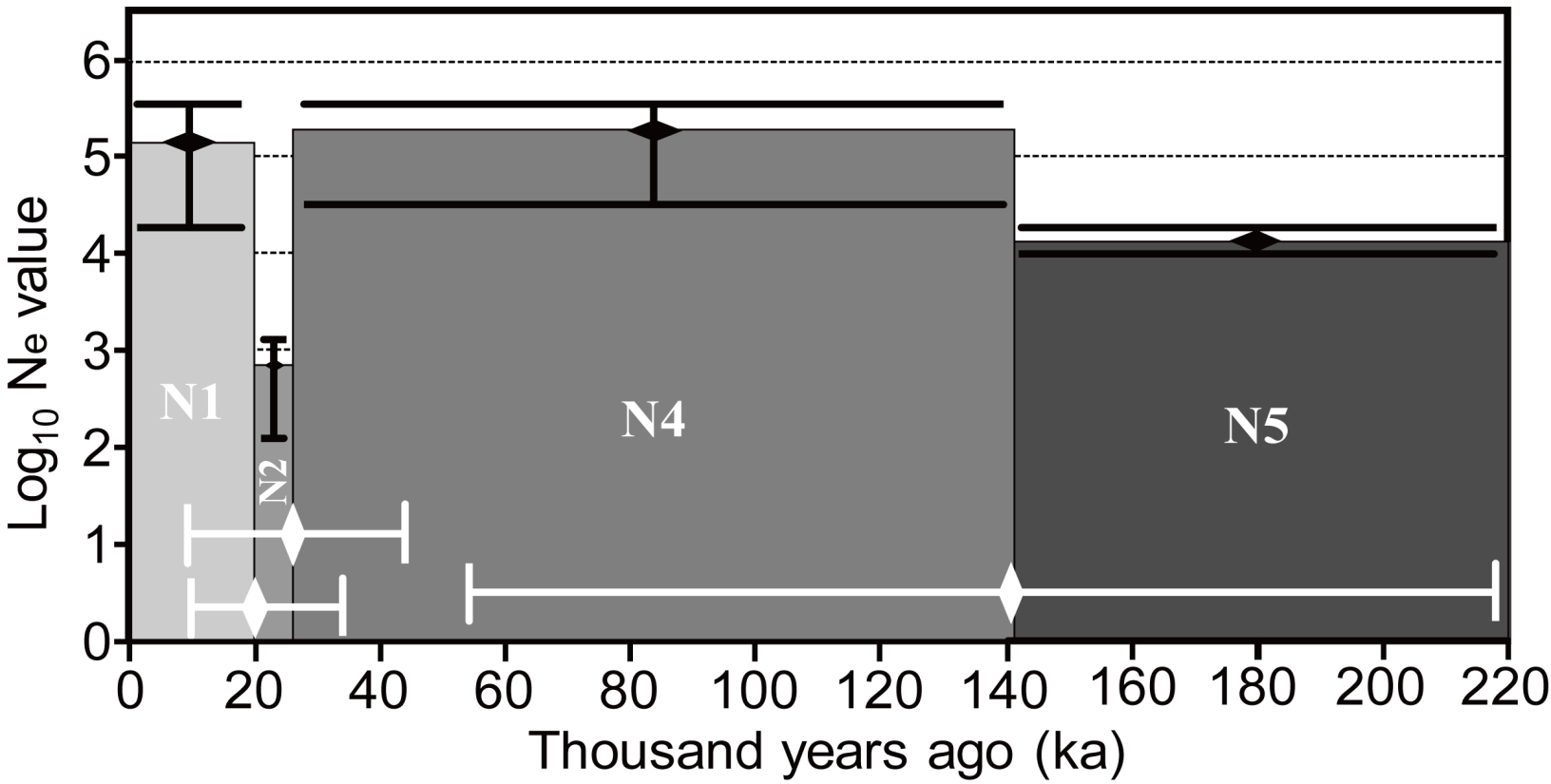

N5
N5
N1
N5
N1
N5
N2
N1
N5
N3
N1
N5
N4
db
N2
N1
Past
t2
t1
Present
Scenario 1
Scenario 2
Scenario 3
Scenario 4
Scenario 5
Scenario 6

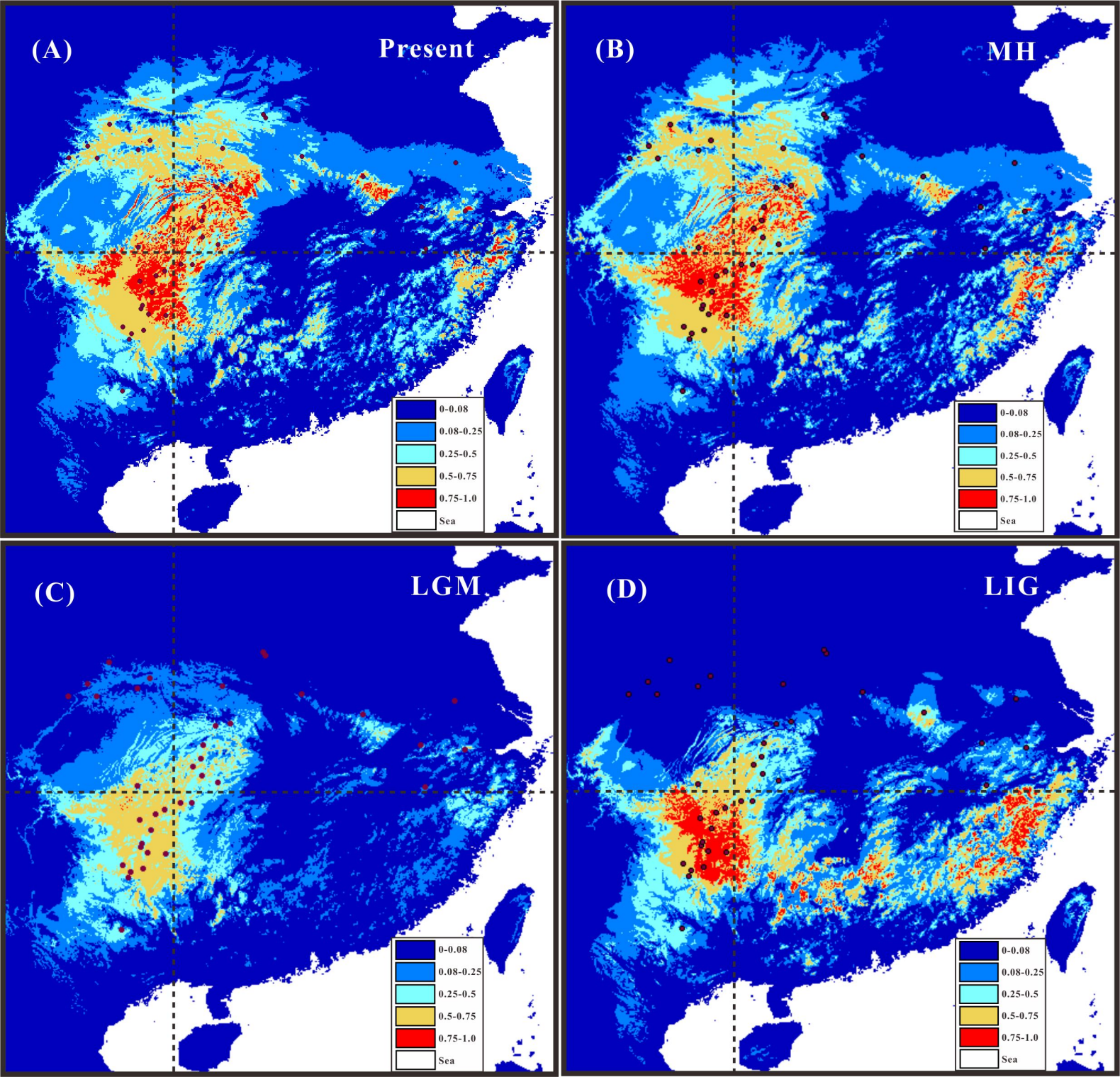

E102
E108
E114
E120
N36
N32
N28
N24
N20
N
Qinling Mts.
Daba Mts.
Yungui Plateau
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
H1
H2
H3
H4
H5
H6
H7
H8
H9
H10
H11
H12
H13
38
Outgroup
>3000
3000
2000
1000
500
300
0
Sea


(A)
<2.3
>=2.3<2.6
>=2.6<2.8
>=2.8<2.9
>=2.9<3.1
>=3.1
0 200 400 km


(B)
<0.34
>=0.34<0.372
>=0.372<0.403
>=0.403<0.434
>=0.434<0.485
>=0.485
0 200 400 km


(A)
K=2
(B)
K=4
(C)
K=6
(D)
Qinling Mts.
Daba Mts.
Yungui Plateau
Cluster1
Cluster2
Cluster3
Cluster4
Cluster5
Cluster6
>3000
3000
2000
1000
500
300
0
Sea
I
II
III
IV
V
E102
E108
E114
E120
N36
N32
N28
N24
N20


6
5
4
3
2
1
0
Log10 Ne value
N1
N2
N4
N5
0
20
40
60
80
100
120
140
160
180
200
220
Thousand years ago (ka)


(A)
Present
0-0.08
0.08-0.25
0.25-0.5
0.5-0.75
0.75-1.0
Sea
(B)
MH
0-0.08
0.08-0.25
0.25-0.5
0.5-0.75
0.75-1.0
Sea
(C)
LGM
0-0.08
0.08-0.25
0.25-0.5
0.5-0.75
0.75-1.0
Sea
(D)
LIG
0-0.08
0.08-0.25
0.25-0.5
0.5-0.75
0.75-1.0
Sea