**ARTICLE**

# Evolutionary history of a beautiful damselfly, *Matrona basilaris*, revealed by phylogeographic analyses: the first study of an odonate species in mainland China

Junli Xue[1] · Haiguang Zhang[1] · Xin Ning[1] · Wenjun Bu[1] · Xin Yu[2,1]

## Abstract

*Matrona basilaris* Selys, 1853 is a damselfly distributed mainly in mainland China. A total of 423 individuals from 48 populations covering almost the entire range were sampled to explore the genetic diversity, phylogeographic structure, and demographic dynamics of the species using sequences of three mitochondrial genes (COI, COII, and ND1) and a nuclear (ITS1 + 5.8 S + ITS2) gene. Phylogenetic tree, median-joining network, and BAPS analyses indicated a four-group division of the entire population, and the divergence event was estimated to have occurred in the middle Pleistocene. The diverse terrain of mainland China as well as past climatic oscillations were assumed to have shaped the current phylogeographic pattern of *M. basilaris*. Multiple lines of evidence supported population expansion in Group 1 and Group 2 but not in Group 3 or Group 4. The expansion times corresponded to the transition phase from the LIG (~0.14–0.12 Mya) to the LGM (~0.021–0.018 Mya). The pre-LGM expansion model reflected a different pattern affecting the historical dynamics of the population of East Asian species caused by Pleistocene climatic changes. Interestingly, Group 2 exhibited a disjunctive distribution pattern. The possible reasons were introgression caused by female-biased dispersal or human phoresy during construction of the Forbidden City during the Ming Dynasty of China.

## Introduction

The global climate has fluctuated during the past two million years, leading to a series of major ice ages during the Quaternary Period. The severe climatic oscillations resulted in drastic environmental changes that profoundly shaped the current distribution and genetic structure of many plants and animals (Hewitt 2000). These events have been well documented in some temperate taxa from Europe and North America, where there were repeated glacial advances and retreats (Coope 1994; Hewitt 1999; Avise 2000; Hewitt 2000, 2004). Plant and animal populations generally exhibited expansion after the Last Glacial Maximum (LGM, ~0.021–0.018 Mya). However, climatic shifts during the Pleistocene varied in magnitude and intensity due to regional differences (Song et al. 2016). In contrast to Europe and North America, East Asia was relatively mildly influenced by glacial cycles and likely possessed a variety of suitable habitats for many species (Porretta et al. 2012; Ye et al. 2014).

Phylogeographic analysis is a powerful method of obtaining insights into the historical processes that have shaped temporal distribution and genetic variation of the species (Avise 2000). Recent phylogeographic research on widely distributed species across East Asia has indicated different responses to historical climatic changes in space and time (Song et al. 2016). Several studies have shown that some species underwent demographic expansion before the arrival of the LGM; such species include plants (Tian et al. 2010; Zhang et al. 2016b), the tree frog *Hyla sarda* (Bisconti et al. 2011), the montane birds *Parus monticolus* (Wang et al. 2013), *Parus major* (Zhao et al. 2012), *Garrulax elliotii* (Qu et al. 2011), and *Leucodioptron canorum*

These authors contributed equally: Junli Xue, Xin Yu

✉ Xin Yu
lannysummer@163.com

[1] Institute of Entomology, College of Life Sciences, Nankai University, Tianjin 300071, China

[2] College of Life Sciences, Chongqing Normal University, Chongqing 401331, China
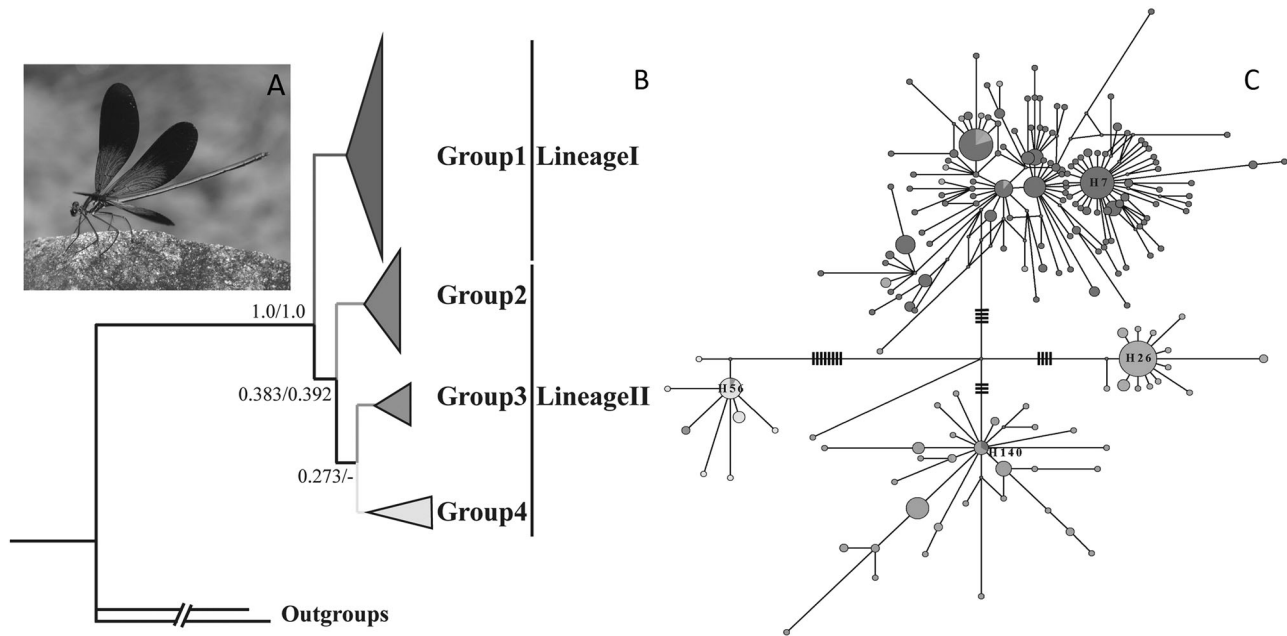
Fig. 1 Male *M. basilaris*, median joining network and neighbor-joining tree of haplotypes. a Ecological photograph of male *M. basilaris*. b Phylogenetic analysis based on mtDNA + ITS. Values besides the nodes indicate bootstrap support (NJ/ML) for major clades. c Median joining network based on mtDNA. Each circle represents a haplotype, and the size of the circle is proportional to the haplotype frequency. The number of the short ligation lines represents mutational steps separating each group. Different colors correspond to different groups defined by BAPS

*canorum* (Li et al. 2009), and some insects, including the semi-aquatic bug *Microvelia douglasi douglasi* (Ye et al. 2014), the cotton pest *Adelphocoris suturalis* (Zhang et al. 2015), and the global invader *Halyomorpha halys* (Zhu et al. 2016). This pattern has been attributed to either the mild Pleistocene climate of East Asia or the intrinsic physiology of the species (e.g., cold-adapted species). Overall, the pre-LGM expansion was opposite to the previous traditional conclusions drawn from temperate species in Europe and North America.

Mainland China, a vast geographic area with complex topography and various climate types, is a main component of East Asia. Many phylogeographic studies have been conducted in mainland China, including studies of plants (Dai et al. 2011), fishes (Yang et al. 2016), birds (Huang et al. 2010; Shi et al. 2014; Qu et al. 2015), and insects (Cheng et al. 2016; Ye et al. 2016; Fang et al. 2017). Phylogeographic patterns revealed in these studies are diverse. Some species were split into western and eastern ranges (Shi et al. 2014); some species were divided into southern and northern ranges (Ye et al. 2014; Cheng et al. 2016; Yang et al. 2016; Ye et al. 2016); and some species were divided into central and peripheral groups (Zhang et al. 2015), while others showed little to no obvious phylogeographic structure (Dai et al. 2011; Wei et al. 2013; Qu et al. 2015; Zhang et al. 2016a; Zhu et al. 2016).

Odonata (damselflies and dragonflies), an ancient insect order, are extremely sensitive to the alteration of their habitats (Khan 2015), which makes them good "thermometers" of environment quality (Silva et al. 2010; Kietzka et al. 2016) and therefore good study subjects for phylogeography (Cordoba-Aguilar and Cordero-Rivera 2005). A number of phylogeographic studies concerning Odonata have been published (Kahilainen et al. 2014; Monroe and Britten 2014; Inomata et al. 2015; Jones and Jordan 2015; Swaegers et al. 2015; Ware et al. 2014), some related to island species of China (Lee and Lin 2012; Xue et al. 2017). However, until now, no comparable research related to odonate species in mainland China has been conducted. *Matrona basilaris*, a beautiful calopterygid damselfly with a metallic-colored body and gorgeous wings (Fig. 1a), is mainly distributed in mainland China; the species also extends into Northern Vietnam and Eastern Laos (*c.f.* Yu et al. 2015). *M. basilaris* prefers to live in cool running water of montane streams or rivers with dense riparian vegetation; the species is thus commonly found at modest altitude (Supplementary Figure S1, Supplementary text). However, *M. basilaris* has an excellent adaptive capacity relative to other species in the genus, and thus displays a large distribution range. We have been focusing on this species since 2007, and we have found it to be a good representative of East Asian Odonata for phylogeographic
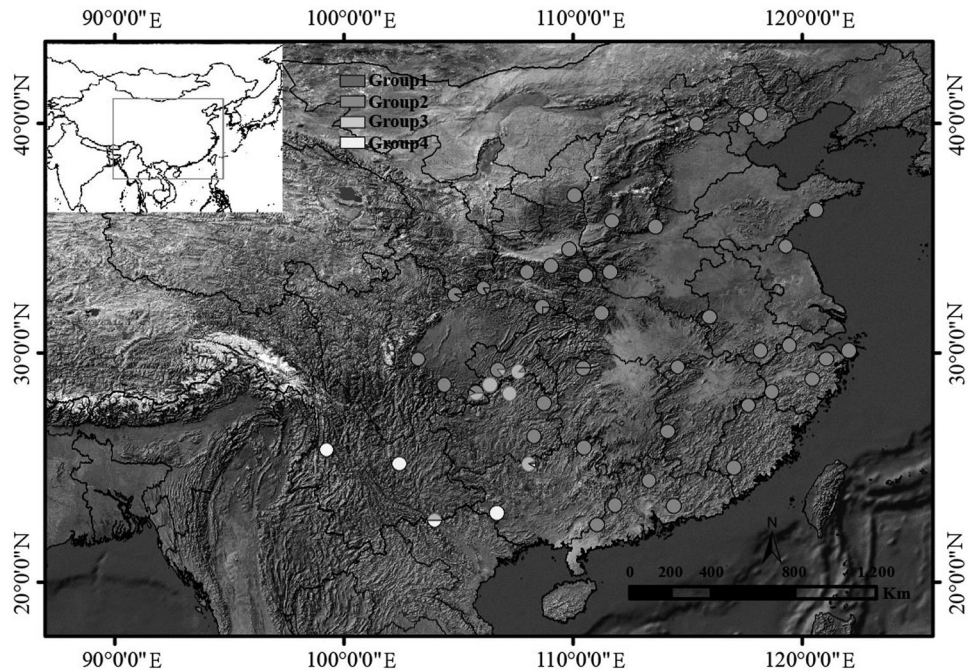
analysis. The present study is based on the largest sample size and the most complete distribution records available. We employed mitochondrial DNA (COI, COII, and ND1) and the nuclear internal transcribed spacer ITS (ITS1-5.8 s rDNA-ITS2) data to carry out a phylogeographic study of *M. basilaris*. The goals were to: (1) explore the genetic diversity and phylogeographic structure, (2) reveal the time and mode of lineage divergence, and (3) investigate the population demographic patterns.

## Materials and Methods

### Sample collection

A total of 423 individuals of *M. basilaris* from 59 locations were collected (Fig. 2; Supplementary Table S1). The sampling locations were finally condensed into 48 populations with a minimum interpopulation distance of 60 km. The sampling scope covered nearly the entire range of this species. Locations were defined according to the combined aspects of geography, environment, and climate. The specimens were immersed in 95% ethyl alcohol in the field and then stored at −20 °C until DNA extraction. All of the samples were identified following Yu et al. (2015). The voucher specimens were deposited in the Institute of Entomology, Nankai University (Tianjin, China).

### DNA extraction and PCR amplification

Complete genomic DNA was extracted using the protocol of the UniversalGen DNA Kit (Beijing ComWin

Biotech Co., Ltd.) according to the manufacturers' instructions. Specifically, one or two legs (instead of the thoracic muscle) were dissected in order to reduce damage to the specimen. Three mitochondrial (COI, COII, and ND1) and two nuclear (ITS and PRMT) genes were amplified and sequenced. Polymerase chain reactions (PCRs) were performed in 40 µl reaction mixtures that each contained 15.5 µl of distilled deionized water, 20 µl of 2× Es Taq MasterMix (Beijing ComWin Biotech Co., Ltd.), 1.5 µl of DNA template, and 1.5 µl of each primer (10 mmol/L). Nucleotide sequences of all relevant primers are summarized in Supplementary Table S2. Reactions were performed in a Biometra T*gradient* Thermocycler (Biometra, Germany) with an initial denaturation at 94 °C for 2 min followed by 35 cycles of denaturation at 92 °C for 30 sec, annealing at 48-56 °C for 30 sec, and extension at 72 °C for 50 sec, with a final elongation step at 72 °C for 8 min. Amplifications were separated on 1.0% agarose gels for quality checking and then sent to commercial companies (BGI Tech Solutions or GENEWIZ Co. Ltd., Beijing, China) for direct sequencing using the same primers as in the PCRs. Fragments that failed in direct sequencing were cloned using a TA-cloning kit by inserting PCR products into the plasmid vector *pEASY*-T3 and transforming into *Trans*1-T1 Phage Resistant Chemically Competent Cells (TransGen Biotech, Beijing, China). At least six single colonies were chosen for colony PCR for each individual, and those containing amplicons of the target size were sequenced using the universal primers M13F and M13R.

## Sequence analyses

Original sequences were assembled and edited manually in BIOEDIT v7.2.3 (Hall 1999). Nucleotide alignments were produced using MAFFT v7.222 (Katoh et al. 2002; Katoh and Standley 2013) and then trimmed to uniform lengths under default settings in BIOEDIT. All of the variable positions were confirmed by visual inspection, and only unambiguous peaks were considered. Alignments of protein coding genes were translated to amino acids to detect frameshift mutations and premature stop codons, which may indicate the presence of nuclear pseudogenes. The mtDNA genes were concatenated as a whole dataset for the subsequent population analyses. Numbers of variable and parsimony informative sites were counted using MEGA v5.0 (Tamura et al. 2011).

## Genetic diversity and structure

Population genetic parameters including the number of haplotypes ($N_h$), haplotype diversity (Hd) and nucleotide diversity ($\pi$) for each population, group (result of Bayesian analysis of population structure below), and overall population were calculated using DNASP v5.0 (Librado and Rozas 2009). Genetic diversity measures for each population were visualized through spatial analysis in ARCGIS v10.0 (ESRI, CA, USA). The population differentiation statistics $G_{ST}$ and $N_{ST}$ were estimated using the program PERMUT v2.0 (Pons and Petitt 1996; Burban et al. 1999). Generally, we considered that phylogeographic structure was evidenced by $N_{ST}$ being significantly larger than $G_{ST}$ using a permutation test with 1000 permutations (Pons and Petitt 1996). Genetic groups were defined by Bayesian analysis of population structure in BAPS v6.0 (Corander et al. 2008). BAPS adopted a stochastic optimization algorithm and applied a Bayesian model to predict the likelihood of a population structure, which could then be used for comparison of different clustering solutions. The "spatial clustering of groups" module was used in the mixture analysis on the basis of the defined 48 populations, and values of $K$ from 2 to 10 were tested. Differences among groups were tested by a hierarchical analysis of molecular variance (AMOVA) (Excoffier et al. 1992) using $F$-statistics. The significance of fixation indices was tested with 1000 random permutations. $F_{ST}$ was estimated in ARLEQUIN v3.5 (Excoffier and Lischer 2010) to quantify the differentiation between pairwise populations and pairwise genetic groups, with statistical significance tested by 1000 non-parametric permutations at the 5% significance level. The significance of isolation by distance among populations was tested using a Mantel test with 1000 randomizations of matrices of pairwise genetic ($F_{ST}$) and geographic distances. The analysis was performed in IBDWS (http://ibdws.sdsu.edu/~ibdws/distances.html).

## Phylogenetic constructions and networks

All of the sequences obtained were collapsed into haplotypes and saved in different formats through DNASP v5.0 (Librado and Rozas 2009) for different phylogenetic analyses. *M. oreads* and *Vestalaria venusta* were designated as outgroups. Phylogenetic topologies were constructed based on mtDNA, nuDNA and the combined dataset using neighbor-joining (NJ), maximum likelihood (ML), and Bayesian inference (BI) methods. NJ and ML analyses were carried out in MEGA v5.0, whereas BI was performed using MRBAYES v3.1.2 (Huelsenbeck and Ronquist 2001). The Kimura 2-parameter model was adopted in the NJ analysis. The optimal model of nucleotide substitution was explored with MEGA v5.0 for ML analyses. Finally, the HKY $+$ I $+$ G model for mtDNA and the GTR $+$ I $+$ G model for the combined dataset were employed. "Use all sites" was chosen as the treatment for gaps and missing data. The node support values were obtained from 1000 bootstrap replicates for both analyses. For the Bayesian analysis, MRMODELTEST v2.2 (Nylander 2004) was employed to explore the best-fit model based on the Akaike information criterion (AIC). Similar appropriate models for different datasets in the ML analyses were also found for the BI analyses. Two independent runs and Markov Chain Monte Carlo (MCMC) were simultaneously conducted for 20 million generations with a sample frequency of 1000. Convergence between runs was monitored by the average standard deviations of split frequencies (<0.01) and the potential scale reduction factor (PSRF, approaching 1.0) (Gelman and Rubin 1992). The effective sample size (ESS) for each parameter was also checked using TRACER v1.5 (Rambaut and Drummond 2007) to assess the stability (ESSs > 200). The first 25% of the sampled trees were discarded as a burn-in phase. Posterior probabilities (PP) were determined by producing a 50% majority-rule consensus tree after discarding the burn-in generations. All of the trees obtained were viewed and edited in FIGTREE v1.4 (Rambaut 2012). In addition, to visualize relationships among haplotypes, median-joining (MJ) networks were constructed using the software NETWORK v5.0 (Fluxus Technology, Suffolk, UK).

## Gene flow

The mutation-scaled effective population size θ and migration rate M were estimated using the program MIGRATE-N v3.6.11 (Beerli and Palczewski 2010) under Bayesian inference. A series of preliminary runs were performed to explore the prior distributions and run conditions.

One long MCMC chain and five replicates visiting 50 million genealogies at a sampling increment of 100 were run, and the first 10% of the samples were discarded as a burn-in. In addition, to improve the MCMC searches, a static heating scheme with four chains at temperatures of 1.0, 1.5, 3.0, and 1 000 000 was used. MIGRATE-N was independently run three times. The initial θ and M were inferred from $F_{ST}$ values, and parameters generated from the first run were used as starting values for the next run.

## Divergence time estimation

The timing of divergence with the 95% highest posterior density (HPD) between lineages of *M. basilaris* was estimated using a Bayesian discrete phylogeographic approach based on all mtDNA sequences from each population using the BEAST v2.21 software package (Bouckaert et al. 2014). *M. nigripectus* and *M. oreads* from the same genus and *Vestalaria* sp and *Neurobasis chinensis* from related genera were chosen as outgroups. The optimal nucleotide substitution model (HKY + I + G) was estimated using MODELTEST v3.7 (Posada and Crandall 1998) under the AIC. The substitution rate 1.77% per site per million years for the COI gene was used (Papadopoulou et al. 2010). The substitution rate of the combined genes was scaled based on the overall mean p-distance according to the substitution rate of COI (Song et. al. 2016). A relaxed uncorrelated lognormal molecular clock model was initially applied in order to give an indication of how clock-like the data were (measured by the ucld.stdev and coefficient of variation parameters). If the ucld.stdev and coefficient of variation parameters were close to 0, then the data were considered as clock-like, and the relatively simple strict clock model was used. Otherwise, the relaxed clock model was employed. A coalescent constant population model was used as the tree prior model. MCMC simulations were run for 100 million generations sampling every 1000 steps. Convergence of the analysis was checked using TRACER to make sure the ESSs were larger than 200. A maximum clade credibility (MCC) tree was summarized in TREEANNOTATOR v2.2.1 (part of the BEAST package) with "mean height" after discarding the first 10% as a burn-in period. Finally, the MCC tree was viewed and modified using FIGTREE.

## Demographic history

Tajima's *D* (Tajima 1989) and Fu's *Fs* (Fu 1997) statistics were used to detect departures from the mutation-drift equilibrium. Significant negative values indicate population expansion, bottlenecks, or natural selection (Fu 1997). A pairwise mismatch distribution analysis (Slatkin and Hudsont 1991), which considers the distribution of pairwise differences between haplotypes, was conducted with 1000 permutations. The 'goodness-of-fit' of the observed and expected distributions of pairwise differences was tested using parametric bootstrapping with the sum of squared deviations (SSD) and Harpending's raggedness index (*r*) (Harpending 1994). The above analyses were conducted in ARLEQUIN. In addition, two methods were used to estimate the timing of population demographic expansion. First, this timing was estimated directly based on the statistic τ (tau) of the mismatch distribution expressed in units of mutational time. This estimate was translated into absolute time (*t*) in years using the formula $\tau = 2ut$ under the sudden expansion model (Rogers and Harpending 1992), where *u* is the mutation rate per year of the whole sequence. Here, the mutation rate was the same as that used in the divergence time estimation. Second, a Bayesian skyline plot (BSP) analysis (Drummond et al. 2005), which uses MCMC integration under a coalescent model, was implemented in BEAST to estimate the change in population size over time. The analyses were only performed based on the mtDNA dataset, because no distinct clade was revealed from ITS, and the mutation rate of this gene is unknown for Odonata. At least 10 million generations or more according to the size of each dataset were run, with sampling every 1000 generations under a strict clock assumption. The concatenated sequences were divided into three partitions based on different genes, and the best-fit model was applied for each partition. To ensure that the MCMC converged to its stationary distribution, we monitored ESS values and reconstructed the demographic history through time in TRACER. Neutrality tests, mismatch distribution and BSP analyses were performed for each group that was inferred from the Bayesian analysis of population structure.

## Results

### Genetic diversity

The final mtDNA dataset included 406 sequences, each 2016 bp in length (COI: 955 bp, COII: 602 bp and ND1: 459 bp). No insertions, deletions, or stop codons were found. A total of 201 polymorphic sites were discovered (105 of these were parsimony informative) defining 181 haplotypes. The overall haplotype (Hd) and nucleotide (π) diversities were 0.975 and 0.00487, respectively. The highest haplotype diversity was detected in Group 2, followed by Group 1, Group 3, and Group 4 (Supplementary Table S1). When small populations (n < 5) were excluded, populations with high haplotype diversity were clustered in the Qinling Mountains (SAXFN, SCSM), west Sichuan (SCMY) and eastern area, including the Dabie and Wuyi Mountains (AHHuB, FJWY, ZJLA, ZJQZ) (Supplementary Figure S2). Furthermore, 393 sequences of the ITS gene

**Fig. 3** Tessellation illustration of Bayesian analysis of population structure based on mtDNA + ITS. The graph in the top left corner shows the changes of log (ml) values with different K values
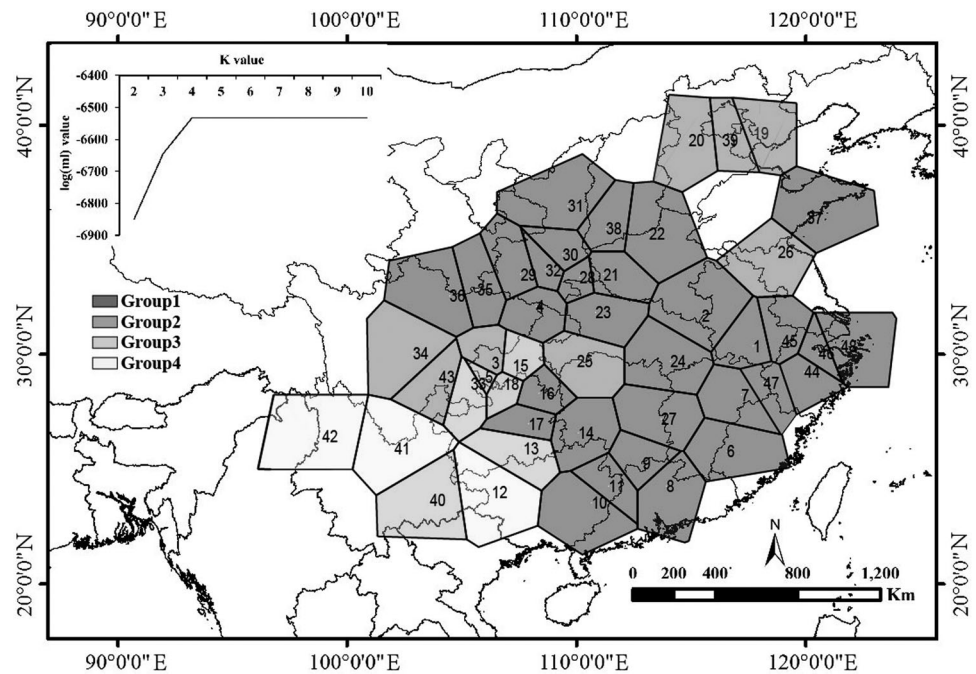


(672 bp) were obtained; three variable sites were detected, and these sites defined three haplotypes. Due to the deficiency of diversity information, the ITS data by itself was only used as a reference in the study (Supplementary Figures S3, S4). One hundred and sixty-nine sequences of the nuclear gene PRMT were obtained from 45 populations. Among these, multiple heterozygous loci were found, and in some cases the number of heterozygous loci numbered up to seven; as this may interfere with the analysis, PRMT was not used in the later analyses.

## Phylogeographic structure

Strong phylogeographic signals were detected on the basis of both mtDNA ($N_{ST}$: 0.565, $G_{ST}$: 0.147, $P < 0.001$) and mtDNA + ITS ($N_{ST}$: 0.585, $G_{ST}$: 0.156, $P < 0.001$) data, i.e., the value of $N_{ST}$ was significantly higher than $G_{ST}$. However, no distinct signal was detected for the ITS data ($N_{ST}$: 0.932, $G_{ST}$: 0.938, P > 0.05). BAPS analysis based on both mtDNA and mtDNA + ITS recovered four groups as the optimal partition of populations (Fig. 3). Group 1, which included 31 populations, was the most widespread, distributed mainly in central and southeast China. Group 2 included eight populations and was separated into western (Sichuan) and northeastern (Hebei and Jiangsu) subgroups. Group 3 included six populations mainly from Guizhou. Group 4 included three populations from Yunnan and Guangxi. AMOVA analysis based on these four groups indicated a significant level of differentiation among groups (58.1% of the variation, $F_{CT} = 0.581$, $P < 0.001$) compared to the level among all populations (31.1%, $F_{SC} = 0.25771$,

$P < 0.001$) or populations within groups (10.8%, $F_{ST} = 0.68898$, $P < 0.001$). This result confirmed the consistency of the group definitions from BAPS. In addition, the pairwise $F_{ST}$ values among the defined groups were large and significant (0.5513~0.6944, $P < 0.001$), suggesting little contemporary gene flow, which could also be implied from the pairwise $F_{ST}$ values among populations (Supplementary Figure S3). Meanwhile, BAPS showed a simple two major group partition based on ITS (Supplementary Figure S4). The Mantel test based on mtDNA and ITS data found a significant but weak positive correlation between genetic divergence and geographic distance (Supplementary Figure S5).

## Phylogenetic analyses and estimation of divergence time

The phylogenetic analyses using NJ, BI and ML methods based on mtDNA and mtDNA + ITS all generated similar results, i.e., four separate branches nested into two distinct geographically structured lineages (Fig. 1b). Lineage I contained only haplotypes from Group 1, whereas Lineage II was comprised of the remaining three groups, although the within-group support values were relatively low. This topology corresponded to the result recovered by BAPS (Fig. 3). The median-joining network based on the mtDNA dataset showed a population structure similar to that retrieved in the phylogenetic analyses, with four haplogroups distinguished (Fig. 1c). A star-like radiation in Group 1 was indicated, with the dominant haplotype (H7) lying in the center. A similar case was observed in Group 2,

| Group | $D$ | $F_S$ | SSD | r | $\tau$ | Expansion time(Ka) |
|---|---|---|---|---|---|---|
| Group 1 | −2.408*** | −24.975*** | 0.0004 ns | 0.0058 ns | 4.75 [2.94,9.80] | 79.9 [49.4,164.8] |
| Group 2 | −1.917*** | −18.127*** | 0.0063 ns | 0.0081 ns | 3.23 [1.39,16.49] | 54.3 [23.4,277.3] |
| Group 3 | −1.602** | −4.838 | 0.0423 ns | 0.0289 ns | — | — |
| Group 4 | −2.129*** | −2.085 | 0.0036 ns | 0.0317 ns | — | — |

Results for Tajima's $D$, Fu's $F_S$, SSD, Harpending's raggedness index ($r$), and mutation time parameter $\tau$ calculated from a sudden expansion demographic model. Brackets contain 95% confidence intervals

*ns* not significant

\*\*\*$P < 0.001$, \*\*$P < 0.01$

with H140 in the center. Both Group 3 and Group 4 had their own dominant haplotypes, and the derivatives were connected by only 1–2 mutational steps. The phylogenetic topology based on ITS indicated an unparsed state, and the network showed a very simple structure (Supplementary Figure S6). The substitution rate of the combined mtDNA sequences was estimated as 1.475% substitutions per site per million years. Comparison of the results indicated that a strict molecular clock model performed better than an uncorrelated lognormal relaxed molecular clock model, with the ucld.stdev <0.1. The Bayesian tree obtained using BEAST showed the same topology as the phylogenetic tree discussed above (Supplementary Figure S7). The initial divergence of two main lineages was dated to approximately 0.41 Mya (95% HPD: 0.27–0.46 Mya). The split within Lineage II occurred soon after the initial isolation, with Group 2 first branching off at about 0.35 Mya (0.23–0.40 Mya).

### Gene flow and demographic history

Results of three independent runs in MIGRATE-N based on combined mtDNA were congruent, and the average value was used for interpretation. Gene flow among groups was asymmetric; the predominant direction of flow was to Group 3 from the other groups, among which the flow from Group 1 was the strongest (Supplementary Figure S8 and Supplementary Table S3). Neutrality tests revealed that major deviations occurred in Group 1 and Group 2, with significant negative Tajima's $D$ and Fu's $F$s values. Group 3 and Group 4 both displayed negative values, but Fu's $F$s was not significant in both cases (Table 1). The observed mismatch distributions for Group 1 and Group 2 presented a unimodal pattern and closely fitted the curve of the simulated distribution, which again appeared to support the above scenarios (Fig. 4a). For Group 3 and Group 4, the mismatch distributions exhibited multimodal shapes, indicating a relatively stable population size (Fig. 4a). According to the values of $\tau$, expansion occurred at about 0.08 Mya for Group 1 and 0.054 Mya for Group 2 (Table 1). The Bayesian skyline plot suggested a long period of constant population size followed by a subsequent

expansion occurring at about 0.075 Mya for Group 1. The expansion trend increased sharply around 0.05 Mya until 0.02 Mya and then reached a plateau. Group 2 also displayed a gentle growth trend since 0.1 Mya. In contrast, analyses for both Group 3 and Group 4 could not reject population stability (Fig. 4b).

## Discussion

### Population divergence and demographic history

According to Yu et al. (2015), *M. nigripectum*, the ancient rare species restricted to the Himalayan and Hengduan areas, is the basal branch of the subgenus *Matrona*. Newly obtained sequence data of *M. nigripectum* in the present study helped us to make a more robust evaluation of the lineage divergence. Multiple lines of evidence, including $N_{ST}$ / $G_{ST}$, BAPS, AMOVA, phylogenetic, and network analyses, all supported a four-group division of the 48 populations of *M. basilaris*. Group 1 was the largest, including 31 populations that were mainly distributed in eastern and central mainland China (Figs. 2, 3). Groups 2–4 had a relatively close relationship (Fig. 1b) and were scattered from the south along the western area toward the north (Figs. 2, 3). Normally, population differentiation events are closely related with factors such as climatic oscillations, geographic barriers, and dispersal ability. The divergence time of *M. basilaris* populations was approximately 0.41 Mya (Supplementary Figure S7), i.e., the period of middle Mid-Pleistocene, corresponding to Marine Isotope Stage (MIS) 12 (0.42–0.48 Mya) (Imbrie et al. 1984; Raymo 1997). During this time, the global climate transformed at a dominant cycle from 41,000 to 100,000 years that has been termed the 'Mid-Pleistocene Revolution' (Wu et al. 2002). The climate of the Asian mainland during this period also underwent a series of cycles oscillating between cold and warm, dry and wet, and continuing until the present (Shi 2002). Furthermore, the diverse topography of mainland China, including features such as the Yunnan-Guizhou Plateau, the Hengduan Mountains, the Qinling Mountains, the Dabie Mountains, and the Wuyi Mountains, will have
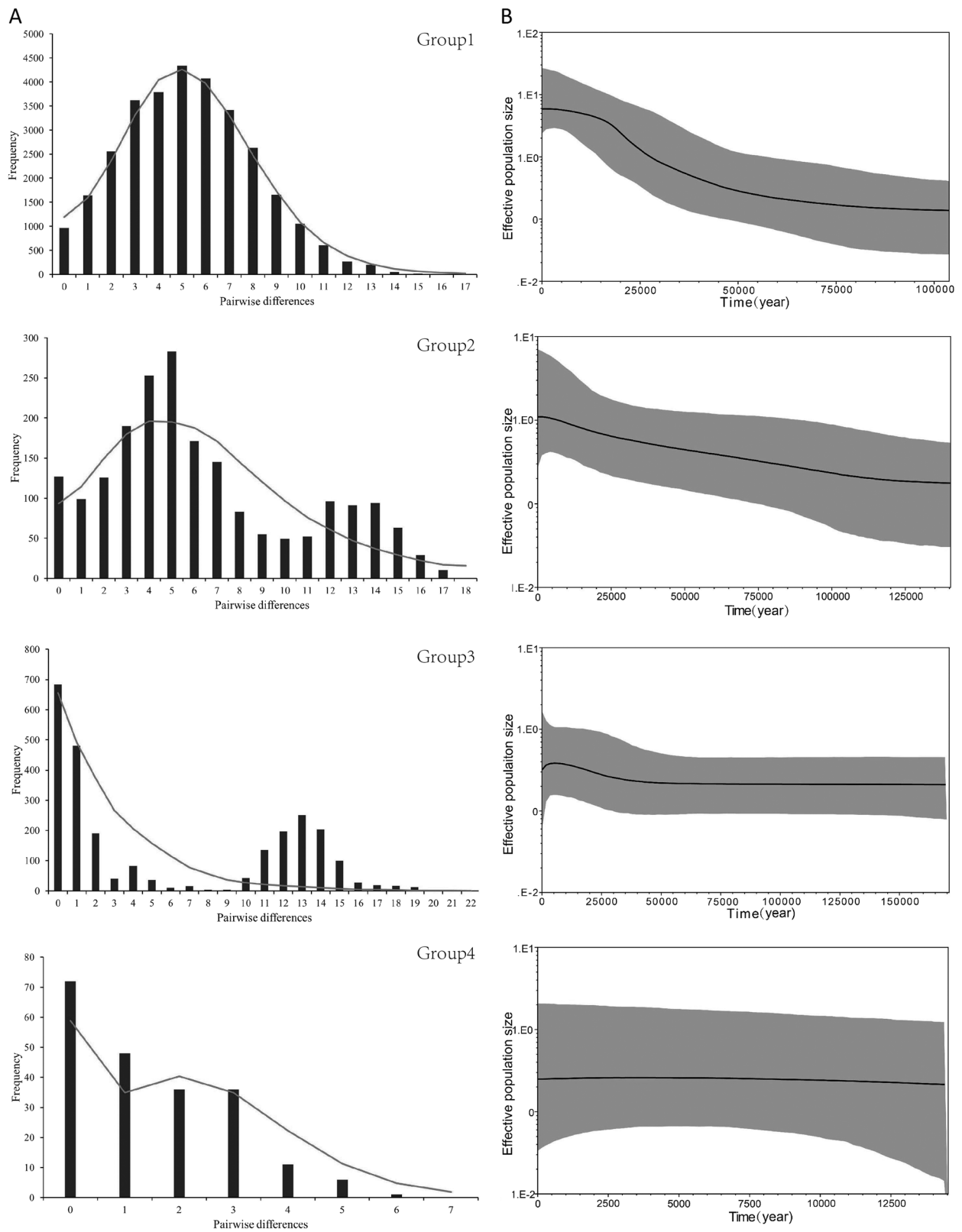
produced both significant geographic barriers and population refugia (Wei et al. 2014; Huang et al. 2015).

Many studies have shown that high or low temperature could drive species distribution ranges toward higher or

◀ **Fig. 4** Pairwise mismatch distributions and BSP for *M. basilaris* inferred from the mtDNA sequences. **a** Pairwise mismatch distributions. The abscissa indicates the number of pairwise differences between compared sequences, and the ordinate is the frequency for each *F* value. Bars represent the observed distribution of pairwise frequencies, while the solid line shows the expected distribution. **b** BSP for *M. basilaris*. The *x*-axis represents time (millions of years ago), and the *y*-axis represents the estimated scaled effective population size. The solid lines indicate the median value of the effective population size. The upper and lower limits of the light blue trend line represent 95% confidence intervals

lower elevations at temperate latitudes (Parmesan 2006; Wilson et al. 2007; Lenoir et al. 2008). As a montane stream dweller, *M. basilaris* is sensitive to temperature changes and is partial to a relatively cool environment, as our own field work experience has shown (Supplementary text, Table S1). We speculated that when the climate was warm and moist, *M. basilaris* preferred to live in relatively high montane streams that were rich in dissolved oxygen; when the climate became cold and dry, the species was driven to lower elevations to track its preferred habitat. If the mountain was large and high enough, for example, the Hengduan Mountains, the damselflies may not have needed to move to the valleys or plains to find water; thus, the species would be restricted to this mountain area without population expansion or gene exchange. This was the case in the groups included in Lineage II, especially for Group 3 and Group 4, each of which had a high frequency of its dominant haplotype (51.4 and 57.1%, respectively) and a relatively low level of genetic diversity (Supplementary Figure S2). Conversely, if the mountain was not so high and large, as in the mountains of Dabie, Daba, and Wuyi, *M. basilaris* may have moved to the plains and expanded its range over quite a large area. This was the case for Lineage I (Group 1) in the eastern area.

Both Group 1 and Group 2 displayed evidence of population expansion starting at 0.075–0.1 Mya, i.e., the last interglacial period (LIG, ~0.14–0.12 Mya); afterward, along with the climate becoming cold and dry in mainland China (Shi 2002), the expansion trend was increasingly intensive until the LGM, perhaps continuing until the present (Fig. 4). This is just the latest example of the scenario suggested above, which is contrary to the general conclusion derived from phylogeographic studies of species in North America and Europe, where the expansion commonly started from refugia after the LGM (Hewitt 2000). This pattern has been seen in a number of recent works focusing on East Asian species (Li et al. 2009; Tian et al. 2010; Bisconti et al. 2011; Dai et al. 2011; Qu et al. 2011; Wang et al. 2013; Ye et al. 2014; Zhang et al. 2015; Zhang et al. 2016b). As stable refugia during in warming stages (providing cool habitats for *M. basilaris*), almost all of the

eastern montane areas showed high genetic diversity (Supplementary Figure S2).

Group 2 is distributed in the Hengduan Mountains, where the altitude is usually high. According to our inference above, it should have shown less expansion. However, this group mainly occurs near an area of the Asian mainland, the Sichuan Basin, that caused major differences. The altitude within the basin is relatively low, and the topography is that of a flat plain. Meanwhile, mountains at the western edge of the Sichuan Basin are higher than those of the eastern edge. All of these factors would have helped *M. basilaris* to expand along the edges of the basin eastward, at times partially crossing the basin. Our results have recovered a gradient of gene exchange of Group 2 with both Group 1 and Group 3 along the edge of the Sichuan Basin at both north and south sides (Fig. 2). Therefore, we inferred that the climatic oscillations and the variations in the terrain in mainland China have led to differentiation and expansion of *M. basilaris* populations.

## Disjunctive distribution of Group 2

It is somewhat puzzling that Group 2 exhibited a disjunctive distribution, i.e., the main part was located in the Sichuan Basin and the surrounding area, whereas the northern part was mainly located in Hebei Province near Beijing (Figs. 2, 3). Herein, we propose two mutually independent hypotheses for the disjunctive distribution: 1) sex-biased dispersal caused introgression, and 2) human phoresy.

Dispersal ability in many species differs between genders, leading to sex-biased dispersal. Beirinckx and Forbes (2006) reported that most damselflies, including *Calopteryx*, the sibling genus of *M. basilaris*, had female-biased dispersal (Dumont et al. 2005). To date, no study has investigated whether *M. basilaris* has sex-biased dispersal. However, according to our field observations, female-biased dispersal exists at least to some degree in this species (Supplementary text). A Mantel test revealed a weak positive correlation between genetic and geographic distances based on mtDNA (Supplementary Figure S6); this may partially confirm the existence of female-biased dispersal (Cooper et al. 2011). Therefore, according to the first hypothesis, if the simple population structure implied by ITS (Supplementary Figure S6) was correct, one can imagine that female-biased dispersal led some of the populations in Group 1 (Fig. 3) to expand northwestward and gradually replace some of the original populations of Group 2. Finally, Group 2 was separated into northern and southern parts, exhibiting a distribution pattern similar to that produced by BAPS based on mtDNA (Fig. 3).

The other hypothesis sounds visionary but can also explain the pattern. It was 600 years ago during the Ming

Dynasty of China when Emperor Yongle decided to construct his imperial city in Beijing (the Forbidden City). This project needed large amounts of a special and rare type of timber, the Nan, which was mainly derived from species in the genera *Phoebe Nees* and *Machilus Nees* in the family Lauraceae. Since these trees were dense in Sichuan (the Hengduan Mountain area), the emperor requisitioned laborers to travel there and transport the timbers to Beijing and surrounding areas. This is a well-known historical fact that has been recorded in the literature of Chinese culture. The logs were usually so huge that they could only be transported via waterways. According to Yun (2006), the main transit lines were from Sichuan to the Yangtze River, and then either through the Beijing-Hangzhou Grand Canal or along the coast to Tianjin and then to Beijing (Supplementary Figure S9). This was very tough work, not only because of the long distance but also because many small dams needed to be built on montane streams where the trees were growing. The aim was to gradually increase the water depth to a level deep enough to float the woods. Therefore, it normally would take 4–5 years to transport the timbers to Beijing (Yun 2006). Females of *M. basilaris* are likely to oviposit on submerged woods or logs in streams. This behavior has been confirmed by many of our field observations (Supplementary Figure S10). Thus, there would have been an opportunity for females from Group 2 (in Sichuan) to lay eggs in timbers transported 600 years ago, thus sending their descendants to Beijing. Compared with the journey in montane areas, it was easier and therefore faster in the Yangtze River, the Beijing-Hangzhou Grand Canal, and on the sea, since the depth of water was no longer relevant. The workers, after leaving the mountains in summer or autumn (the rainy season with enough water), would have tried to transport the timbers to Beijing as soon as possible, since rivers in northern China will freeze in winter. Thus, during the following spring the eggs in the logs may have hatched, and some of the larvae could have adapted to the local climate and become founders of new populations.

Both of the above hypotheses need further research for verification. We intend to conduct studies using more sensitive molecular markers and to complete our samples to investigate whether there is genetic divergence between the north and south parts of Group 2, and if so, to estimate the time of divergence. Thus, we may confirm which of the hypotheses is the true reason for the disjunction, or if both events occurred in nature.

## Data Archiving

Sequence data used in this article must be made available in the journal's database. Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.r4v7623

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

Avise JC (2000) Phylogeography: The History and Formation of Species.. Harvard University Press, Cambridge

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

Beirinckx K, Forbes MR (2006) Sex biases in dispersal and philopatry: insights from a meta-analysis based on capture-mark-recapture studies of damselflies. Oikos 113:539–547

Bisconti R, Canestrelli D, Colangelo P, Nascetti G (2011) Multiple lines of evidence for demographic and range expansion of a temperate species (*Hyla sarda*) during the last glaciation. Mol Ecol 20:5313–5327

Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D et al. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput Biol 10:e1003537

Burban C, Petit RJ, Carcreff E, Jactel H (1999) Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera:Matsucoccidae) in relation to the genetic structure of its host. Mol Ecol 8:1593–1602

Cheng R, Jiang N, Yang X, Xue D, Liu S, Han H (2016) The influence of geological movements on the population differentiation of *Biston panterinaria* (Lepidoptera: Geometridae). J Biogeogr 43:691–702

Coope GR (1994) The response of insect faunas to glacial-interglacial climatic fluctuations. Philos Trans R Soc B 344:19–26

Cooper SJB, Harvey MS, Saint KM, Main BY (2011) Deep phylogeographic structuring of population of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: evidence for long-term refugia within refugia. Mol Ecol 20:3219–3236

Corander J, Marttinen P, Sirén J, Jing T (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinf 9:539

Cordoba-Aguilar A, Cordero-Rivera A (2005) Evolution and ecology of Calopterygidae (Zygoptera: Odonata): status of knowledge and research perspectives. Neotrop Entomol 34:861–879

Dai CY, Zhao N, Wang WJ, Lin CT, Gao B, Yang XJ et al. (2011) Profound climatic effects on two East Asian black-throated tits (Ave: Aegithalidae), revealed by ecological niche models and phylogeographic analysis. PLoS ONE 6:e29329

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22:1185–1192

Dumont HJ, Vanfleteren JR, De Jonckheere JF, Weekers PHH (2005) Phylogenetic relationships, divergence time estimation, and global biogeographic patterns of Calopterygoid damselflies (Odonata, Zygoptera) inferred from ribosomal DNA sequences. Syst Biol 54:347–362

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distance among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Fang F, Chen J, Jiang LY, Chen R, Qiao GX (2017) Biological traits yield divergent phylogeographical patterns between two aphids living on the same host plants. J Biogeogr 44:348–360

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol 66:591–600

Hewitt GM (1999) Post-glacial re-colonization of European biota. Biol J Linn Soc 68:87–112

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–913

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc Lond B Biol Sci 359:183–95. discussion 195

Huang XL, Liu DG, Wang D, Shi XQ, Simon JC (2015) Molecular and Quantitative genetic differentiation in sitobion avenae populations from both sides of the Qinling Mountains. PLoS ONE 10:e0122343

Huang ZH, Liu NF, Liang W, Zhang YY, Liao XJ, Ruan LZ et al. (2010) Phylogeography of Chinese bamboo partridge, Bambusicola thoracica thoracica (Aves: Galliformes) in south China: inference from mitochondrial DNA control-region sequences. Mol Phylogenet Evol 56:273–280

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Imbrie J, Hays JD, Martinson DG, Mcintyre A, Mix AC, Morley JJ et al. (1984). The orbital theory of Pleistocene climate: support from a revised chronology of the marine δ 18O record. In: Berger A, Imbrie J, Hays H, Kukla G, Saltzman B (eds) Milankovitch and climate, part I, Reidel Publishing: Dordrecht. pp 269–305.

Inomata N, Hironaka K, Sawada K, Kuriwada T, Yamahira K (2015) Discrepancy in the degree of population differentiation between color-morph frequencies and neutral genetic loci in the damselfly Ischnura senegalensis in Okinawa Island, Japan. Genetica 143:271–277

Jones BR, Jordan S (2015) Genetic consequences of Pleistocene sea-level change on Hawaiian Megalagrion damselflies. J Hered 106:618–627

Kahilainen A, Keranen I, Kuitunen K, Kotiaho JS, Knott KE (2014) Interspecific interactions influence contrasting spatial genetic structures in two closely related damselfly species. Mol Ecol 23:4976–4988

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780

Katoh K, Misawa K, Kuma KI, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066

Khan MK (2015) Dragonflies and damselflies (Insecta: Odonata) of the northeastern region of Bangladesh with five new additions to the Odonata fauna of Bangladesh. J Threat Taxa 7:7795–7804

Kietzka GJ, Pryke JS, Samways MJ (2016) Aerial adult dragonflies are highly sensitive to in-water conditions across an ancient landscape. Divers Distrib 23:1–13

Lee YH, Lin CP (2012) Pleistocene speciation with and without gene flow in Euphaea damselflies of subtropical and tropical East Asian islands. Mol Ecol 21:3739–3756

Lenoir J, Gégout JC, Marquet PA, Ruffray Pde, Brisse H (2008) A significant upward shift in plant species optimum elevation during the 20th century. Science 320:1768–1771

Li SH, Yeung CK, Feinstein J, Han L, Le MH, Wang CX et al. (2009) Sailing through the Late Pleistocene: unusual historical demography of an East Asian endemic, the Chinese Hwamei (Leucodioptron canorum canorum), during the last glacial period. Mol Ecol 18:622–633

Librado P, Rozas J (2009) DnaSPv5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Monroe EM, Britten HB (2014) Conservation in Hine's sight: the conservation genetics of the federally endangered Hine's emerald dragonfly. Somat hineana J Insect Conserv 18:353–363

Nylander JAA (2004). MrModeltest v2. Program distributed by the author. Evolutionary biology centre, Uppsala University, Uppsala, Sweden.

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Mol Biol Evol 27:1659–1672

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. Annu Rev Ecol Evol S 37:637–669

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

Porretta D, Mastrantonio V, Bellini R, Somboon P, Urbanelli S (2012) Glacial history of a modern invader: phylogeography and species distribution modelling of the Asian tiger mosquito Aedes albopictus. PLoS ONE 7:e44515

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818

Qu YH, Luo X, Zhang RY, Song G, Zou FS, Lei FM (2011) Lineage diversification and historical demography of a montane bird Garrulax elliotii-implications for the Pleistocene evolutionary history of the eastern Himalayas. BMC Evol Biol 11:174

Qu YH, Song G, Gao B, Quan Q, Ericson PGP, Lei FM et al. (2015) The influence of geological events on the endemism of East Asian birds studied through comparative phylogeography. J Biogeogr 42:179–192

Rambaut A (2012). FigTree v.1.4. Available at http://tree.bio.ed.ac.uk/software/figtree. Accessed 18 Dec 2012.

Rambaut A, Drummond AJ (2007). Tracer v1.5 Available at: http://tree.bio.ed.ac.uk/software/tracer/. Accessed 30 Nov 2009.

Raymo ME (1997) The timing of major climate terminations. Paleoceanography 12:577–585

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

Shi MM, Michalski SG, Welk E, Chen XY, Durka W, Carine M (2014) Phylogeography of a widespread Asian subtropical tree: genetic east-west differentiation and climate envelope modelling suggest multiple glacial refugia. J Biogeogr 41:1710–1720

Shi YF (2002) A suggestion to improve the chronology of quaternary glaciations in China. J Glaciol Geocryol 24:687–692

Silva DDP, Marco PD, Resende DC (2010) Adult odonate abundance and community assemblage measures as indicators of stream ecological integrity: A case study. Ecol Indic 10:744–752

Slatkin M, Hudsont RR (1991) Mismatch pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555–562

Song G, Zhang RY, DuBay SG, Qu YH, Dong L, Wang WJ et al. (2016) East Asian allopatry and north Eurasian sympatry in long-tailed tit lineages despite similar population dynamics during the late Pleistocene. Zool Scr 45:115–126

Swaegers J, Mergeay J, St-Martin A, Knijf GD, Larmuseau MHD, Stoks R (2015) Genetic signature of the colonisation dynamics along a coastal expansion front in the damselfly *Coenagrion scitulum*. Ecol Entomol 40:353–361

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Tian S, López-Pujol J, Wang HW, Ge S, Zhang ZY (2010) Molecular evidence for glacial expansion and interglacial retreat during Quaternary climatic changes in a montane temperate pine (*Pinus kwangtungensis* Chun ex Tsiang) in southern China. Plant Syst Evol 284:219–229

Wang WJ, McKay BD, Dai CY, Zhao N, Zhang RY, Qu YH et al. (2013) Glacial expansion and diversification of an East Asian montane bird, the green-backed tit (*Parus monticolus*). J Biogeogr 40:1156–1169

Ware JL, Beatty CD, Herrera MS, Valley S, Johnson J, Kerst C et al. (2014) The petaltail dragonflies (Odonata: Petaluridae): Mesozoic habitat specialists that survive to the modern day. J Biogeogr 41:1291–1300

Wei CH, Lu J, Xu LY, Liu G, Wang ZG, Zhao FP et al. (2014) Genetic structure of Chinese Indigenous goats and the special geographical structure in the Southwest China as a geographic barrier driving the fragmentation of a large population. PLoS ONE 9:e94435

Wei SJ, Shi BC, Gong YJ, Jin GH, Chen XX, Meng XF (2013) Genetic structure and demographic history reveal migration of the diamondback moth *Plutella xylostella* (Lepidoptera: Plutellidae) from the southern to northern regions of China. PLoS ONE 8: e59654

Wilson RJ, Gutiérrez D, Gutiérrez J, Monserrat VJ (2007) An elevational shift in butterfly species richness and composition accompanying recent climate change. Glob Change Biol 13:1873–1887

Wu GJ, Pan BT, Guan QY, Gao HS (2002) Review of studies on the Mid-Pleistocene climatic transition and the 100Ka cycle. Adv earth Sci 17:605–611

Xue JL, Yu X, Zhang HG, Chen X, Bu WJ (2017) Population genetics and ecological niche modeling shed light on conservation of the island endemic damselfly *Pseudolestes mirabilis* (Odonata, Pseudolestidae). Hydrobiologia 790:273–286

Yang JQ, Hsu KC, Liu ZZ, Su LW, Kuo PH, Tang WQ et al. (2016) The population history of *Garra orientalis* (Teleostei: Cyprinidae) using mitochondrial DNA and microsatellite data with approximate Bayesian computation. BMC Evol Biol 16:73

Ye Z, Zhu GP, Chen PP, Zhang DL, Bu WJ (2014) Molecular data and ecological niche modelling reveal the Pleistocene history of a semi-aquatic bug (*Microvelia douglasi douglasi*) in East Asia. Mol Ecol 23:3080–3096

Ye Z, Zhu GP, Damgaard J, Chen X, Chen PP, Bu WJ (2016) Phylogeography of a semi-aquatic bug, *Microvelia horvathi* (Hemiptera: Veliidae): an evaluation of historical, geographical and ecological factors. Sci Rep 6:21932

Yu X, Xue JL, Hämäläinen M, Liu Y, Bu WJ (2015) A revised classification of the genus *Matrona* Selys, 1853 using molecular and morphological methods (Odonata: Calopterygidae). Zool J Linn Soc 174:473–486

Yun Y (2006) Record and narrate of the logging project in the constructing of the Forbidden City. Dong Yue Trib 27:167–172

Zhang DL, Zhen Y, Yamada K, Zhen YH, Zheng CG, Bu WJ (2016a) Pleistocene sea level fluctuation and host plant habitat requirement influenced the historical phylogeography of the invasive species *Amphiareus obscuriceps* (Hemiptera: Anthocoridae) in its native range. BMC Evol Biol 16:174

Zhang LJ, Li H, Li SJ, Zhang AB, Kou F, Xun HZ et al. (2015) Phylogeographic structure of cotton pest *Adelphocoris suturalis* (Hemiptera: Miridae): strong subdivision in China inferred from mtDNA and rDNA ITS markers. Sci Rep 5:14009

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

Zhao N, Dai CY, Wang WJ, Zhang RY, Qu YH, Song G et al. (2012) Pleistocene climate changes shaped the divergence and demography of Asian populations of the great tit *Parus major*: evidence from phylogeographic analysis and ecological niche models. J Avian Biol 43:297–310

Zhu GP, Ye Z, Du J, Zhang DL, Zhen YH, Zheng CG et al. (2016) Range wide molecular data and niche modeling revealed the Pleistocene history of a global invader (*Halyomorpha halys*). Sci Rep 6:23192