Impact of Orogeny and Environmental Change on Genetic Divergence and Demographic History of Dipus sagitta *(Dipodoidea, Dipodinae) since the Pliocene in Inland East Asia*

Jilong Cheng, Xue Lv, Lin Xia, Deyan Ge, Qian Zhang, Liang Lu & Qisen Yang

Journal of Mammalian Evolution

ISSN 1064-7554
Volume 26
Number 2

J Mammal Evol (2019) 26:253-266
DOI 10.1007/s10914-017-9397-6

Springer

J Mammal Evol (2019) 26:253–266
DOI 10.1007/s10914-017-9397-6

ORIGINAL PAPER

Impact of Orogeny and Environmental Change on Genetic Divergence and Demographic History of *Dipus sagitta* (Dipodoidea, Dipodinae) since the Pliocene in Inland East Asia

Jilong Cheng[1,2] · Xue Lv[1,2] · Lin Xia[1] · Deyan Ge[1] · Qian Zhang[1] · Liang Lu[3] · Qisen Yang[1]

Published online: 14 June 2017

Abstract Inland East Asia encompasses particular landscapes, including discontinuous large deserts isolated by mountains, and these landscapes have been greatly impacted by the uplift of the Qinghai-Tibetan Plateau (QTP) since the Pliocene. However, little research has been performed on the impact of desertification on the evolutionary history of animals in this area. We examined a widespread desert rodent species, *Dipus sagitta*, to better understand the influences of geological events on the evolutionary history and phylogeographic patterns. We sequenced two mitochondrial genes and three nuclear genes from 237 individuals collected from 43 populations across inland East Asia. Phylogenetic, network, intraspecific delimitation, and population structure analyses identified a structured pattern of geographic differentiation with six well-defined evolutionary clades. High mountains (such as the Tianshan Mountains) and important climate demarcation lines (such as the 200 mm isohyet) were inferred as genetic barriers among the six clades by BARRIER analysis. The most recent common ancestor of *D. sagitta* was estimated to have existed in late Miocene, and the first clade was estimated to have diverged at c. 7.57 Ma, whereas the later clade diversified rapidly at approximately 2.56–1.53 Ma during the Pleistocene. Demographic analyses suggested different demographic histories among the distinct clades, and the continuous expansion that occurred during the climate oscillation period appeared to be more closely related to the increasing aridification caused by orogeny rather than climate oscillation.

Keywords *Dipus sagitta* · Ecological divergence · Geographical isolation · Inland East Asia · Orogenic events · Phylogeographic pattern · Regional aridification

Background

Eurasia possesses the largest mid-latitude arid zone in the world. More than half of this arid zone (over 1.5 million km^2) is located in inland East Asia, which has more complex landforms and a more fascinating geological history than Central Asia (Zhang 2012). This biogeographic area is limited to the west by the Pamir region and the Tianshan Mountains, to the north by the Siberian Taiga, to the east by the Northeast China Plain, and to the south by the Qinghai-Tibetan Plateau (QTP) and the Loess Plateau. In such an area, aridification tends to form strong barriers to the dispersal of organisms that are not adapted to arid environments (Song et al. 2016) but provides new ecological opportunities and promotes diversification of desert organisms (Zhang et al. 2013). Inland East Asia was strongly influenced by aridification, related to long-term global cooling and drying trends since the late Miocene (Miao et al. 2012; Meng et al. 2015), and the tectonic uplift of the QTP, which began in the late Pliocene (Zhang et al. 2008; Song et al. 2013). Furthermore, climatic oscillations

Electronic supplementary material The online version of this article (doi:10.1007/s10914-017-9397-6) contains supplementary material, which is available to authorized users.

✉ Qisen Yang
yangqs@ioz.ac.cn

[1] Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

[2] University of Chinese Academy of Sciences, Shijingshan District, Beijing 100049, China

[3] State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Centre for Disease Control and Prevention, Beijing 102206, China

during the Quaternary induced repeated succession of desert-steppe-forest ecosystems (Li et al. 2002; Zhang 2012) and a substrate shift between sandy desert and stony desert on the Mongolian Plateau (MP) and in adjacent regions.

Topography and climate have shaped the phylogeographic patterns of current populations, separately or together (Maldonado-Sanchez et al. 2016). The geographical isolation of populations is generally caused by complex topographical landforms that limit the gene flow among populations and provide opportunities for intraspecific diversification or even speciation due to genetic drift and natural selection (Qu et al. 2015). For psammophilic animals, high mountains and large water areas between deserts often act as geographic barriers that promote vicariant speciation or population differentiation (Castoe et al. 2007; Phillipsen and Metcalf 2009; Su and Zhang 2013). Moreover, deserts and arid regions generally harbor a relatively high rate of endemism and act as local hotspots of biodiversity because of adaptation to extreme environments and sharp ecological gradients (Durant et al. 2012; Murphy et al. 2012; Wilson and Pitts 2012; Brito et al. 2014). Hence, the repeatedly alternating environments caused by climatic oscillations may have induced ecological divergence between populations as they adapted to differing environments, thereby forming new subdivisions or even new species (Ge et al. 2005; Lei et al. 2014).

Compared with those in inland East Asia, the impacts of geological events and environmental changes on the phylogeographic patterns of species in North America and Europe at the same latitude have been well studied. In East Asia, researchers have primarily focused on the montane ecosystems of southwest China. Moreover, although the arid area in inland East Asia has a unique geological history and natural climate, as well as different genetic patterns (Chen et al. 2008; Song et al. 2013; Lv et al. 2016), few studies have focused on species with ranges that span the entirety of this region.

Small rodents feature short generation times, rapid mitochondrial DNA (mtDNA) evolutionary rates, relatively limited dispersal abilities, and strong associations with specific habitats (Ben Faleh et al. 2012b). The northern three-toed jerboa, *Dipus sagitta*, the only species that has thrived in all the deserts in inland East Asia, is a suitable candidate for inferring the influence of historical geological events and climatic changes on phylogeographic patterns of this area. *Dipus sagitta* is distributed from the Don River in southeastern Europe through Central Asia to Northeastern China and is the only present-day species of its genus *Dipus* (Fig. 1). The distribution map of the species presented herein was slightly modified from that of the IUCN (Batsaikhan et al. 2016). *Dipus sagitta* is mainly specialized for sandy deserts and prefers to inhabit gradual slopes among large sand dunes covered by sparse shrubs and herbs (Shenbrot et al. 2008). Based on the structure of its hind limbs as well as its genitalia and karyotype, *D. sagitta* is divided into two groups, the "*Sagitta*" group and the "*Lagopus*" group (Shenbrot 1991a, b). The "*Sagitta*" group is predominant in our study area. The diploid complement of *D. sagitta* consists of 48 chromosomes (2n = 48), and the chromosomal complement is characterized by the number of autosomal arms, which is 90.

A recent study has investigated phylogeographic patterns displayed by *Meriones meridianus*, *D. sagitta*, and *Allactaga sibirica* (Liao et al., 2016). However, that study area was limited to the Hexi Corridor, and the sample size was small; consequently, the study was unable to reveal the true pattern of *D. sagitta*. In this study, we increased the sample size of *D. sagitta* to cover the species distribution in the study area of inland East Asia and included nuclear loci to explore the phylogeographic patterns more accurately. We aim to (1) determine the phylogeographic pattern of *D. sagitta* and its demographic history; (2) infer the relationship between the major biogeographic events and the observed patterns of *D. sagitta*; and (3) examine the impact of potential geographical barriers on population differentiation, such as mountains and significant climate differences.

Materials and Methods

Study Area

Inland East Asia features a series of discontinuous deserts isolated by huge mountains and large basins, and exhibits variations in climate parameters (Fig. 1) (Zhang 2012). In the western part of this area, the Kunlun, Tianshan, Altay, and Qilian mountain ranges form a succession of peaks with elevations greater than 3000–4000 m and separate the continent into the Tarim, Junggar, Turpan-Hami, and Qaidam basins. The driest parts of this region receive less than 100 mm of annual precipitation. Further to the east, the rainfall on the MP increases to 200 mm, and the elevation of the mountain ranges decreases to 1000–2000 m. The topographic feature formed by the Yinshan Mountains and Helan Mountains constitutes the major boundary between monsoonal and non-monsoonal regions in East Asia, corresponding to the 200 mm isohyet. The region to the east and southeast is the wettest region, exhibiting 200–400 mm of annual precipitation, and is covered with prairie developed on sandy land.

Sampling, Amplification, and DNA Sequencing

This study was carried out in strict accordance with animal research protocol IOZ-2006 approved by the Animal Care Committee of the Institute of Zoology, Chinese Academy of Sciences (IOZCAS). Northern three-toed jerboas were captured with the permission of the local Protection and Research Centre and the Forestry Administration of Gansu, Inner Mongolia, Jilin, Liaoning, Ningxia, Qinghai, and Xinjiang Provinces. All jerboas were captured in cages and

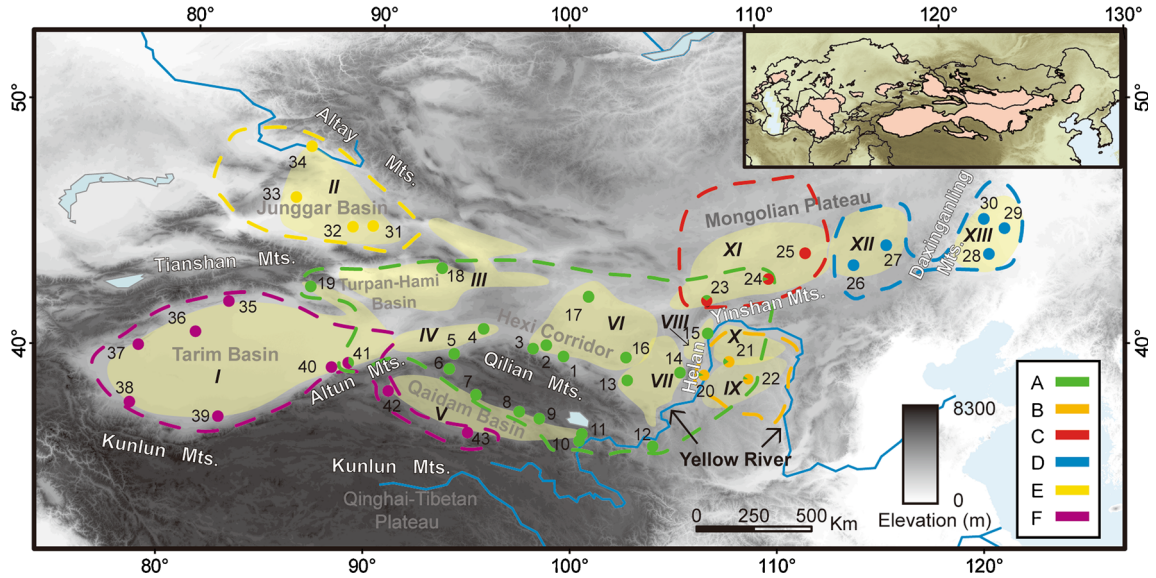

Fig. 1 Sampling sites of the populations of *D. sagitta*, including data from GenBank. The locality codes and coordinates are presented in Supplementary Table 2. The genetic clades A, B, C, D, E, and F are labelled with green, orange, red, blue, yellow, and purple, respectively. The color of each sampling site indicates individuals that belong to the corresponding clade identified by a Bayesian phylogenetic analysis and Bayesian population structure analysis based on mtDNA. Major geographical regions are marked on the map. Large deserts in inland East Asia are indicated by Roman numerals (in italics): *I*, Taklimakan Desert; *II*, Gurbantunggut Desert; *III*, Hashun Gobi Desert; *IV*, Kumtag Desert; *V*, Qaidam Desert; *VI*, Badain Jaran Desert; *VII*, Tengger Desert; *VIII*, Ulan Buh Desert; *IX*, Mu Us Sandy Land; *X*, Kubuqi Desert; *XI*, Gobi Desert; *XII*, Hunshandake Sandy Land; and *XIII*, Horqin Sandy Land. Mts. is short for Mountains

sacrificed via decapitation immediately after capture. All efforts were made to minimize their potential pain and suffering. Intact skulls of all the captured individuals were removed and preserved in 100% ethanol in the field and cleaned after the field survey for further identification. Muscle tissues to be employed for molecular analysis were collected and preserved in 100% ethanol. All specimens and skulls collected in the field were deposited in the mammal collections of the National Zoological Museum, IOZCAS.

A total of 204 individuals of *D. sagitta* were sampled from 35 localities throughout the study area between 2007 and 2015 (Fig. 1). Furthermore, we downloaded 33 additional sequences from eight localities from GenBank. Total genomic DNA was extracted using a TIANamp Genomic DNA Kit (DP304; Tiangen Biotech, Beijing, China) according to the manufacturer's instructions. Two mtDNA sequences [cytochrome *b* (*Cytb*) and cytochrome oxidase subunit I (*COI*)] and three nDNA sequences [exon 10 of the growth hormone receptor (*GHR*), exon 1 of the interstitial retinoid-binding protein (*IRBP*), and a portion of recombination activating gene 1 (*RAG1*)] were amplified via polymerase chain reaction (PCR). All amplifications were conducted in 25 μL reactions containing approximately 25 ng of extracted DNA, 200 μM of each dNTP, 0.2 μM of each primer, 0.75 unit of LA Taq polymerase, and 2.5 μL of 10× buffer. The amplifications were performed using the following PCR profiles: initial denaturation at 94 °C for 3 min, followed by 35 cycles with initial denaturation at 94 °C for 30 s, annealing at 50–60 °C (depending on the primers) for 30 s, and polymerization at 72 °C for 1 min or 2 min (depending on the length of the target); and a final extension at 72 °C for 10 min. The primers and annealing temperatures used, along with the corresponding primary references, are listed in Supplementary Table 1. All sequences were manually inspected and assembled using BIOEDIT 7.2.5 (Hall 1999). Sequence alignments were performed using the MUSCLE algorithm (Edgar 2004) implemented in MEGA 6 (Tamura et al. 2013). All sequences reported in this study were deposited in GenBank (see Supplementary Table 2).

We concatenated the *Cytb* and *COI* genes because they are linked and likely evolved under similar constraints. Population genetic parameters for the total samples and each clade (see Results), such as haplotype diversity (Hd) and nucleotide diversity (Pi) based on combined mtDNA sequence datasets, were calculated to estimate DNA polymorphism using DNASP 5.10.01 (Librado and Rozas 2009). We also calculated the number of polymorphic (segregating) sites and the total number of mutations for each gene.

Phylogenetic Analysis and Divergence Times

To evaluate the intraspecific phylogenetic relationship among the samples, individual gene trees were reconstructed based on the combined mtDNA genes and each nDNA gene using Bayesian inference methods in MRBAYES 3.2.5 (Ronquist et al. 2012). JMODELTEST 2.1.7 (Darriba et al. 2012) was employed to select the best fit model of base pair substitutions for the dataset based on the Akaike information criterion (AIC). Two parallel runs of one cold and three heated

Markov chain Monte Carlo (MCMC) analyses were performed for 20 million generations or more, with trees sampled every 1000 generations to produce convergence (SD < 0.01). The first 25% of the Markov chain samples ($N = 20{,}000$) were discarded as burn-in, and the remaining samples were used to generate majority rule consensus trees. The trees were then edited in FIGTREE 1.4.2 (available at http://tree.bio.ed.ac.uk/software/fig-tree/). To better visualize the relationships among haplotypes, median-joining (MJ) networks (Bandelt et al. 1999) were generated using the combined mtDNA sequences in POPART 1.7 (Leigh and Bryant 2015). Gaps and missing sites in the sequences were excluded from the analysis. Because only *Cytb* gene sequences were obtained from GenBank, we performed the analysis using only the samples collected in the field. The pairwise genetic distances of the mtDNA sequences were calculated with MEGA version 6.04 (Tamura et al. 2013) using the Kimura two-parameter (K2P) distance between and within each clade.

To estimate the divergence times of all identified haplotypes, we applied a Bayesian MCMC analysis implemented in the program BEAST 1.8.2 (Drummond et al. 2012). The analysis was performed using the combined mtDNA data, as mtDNA possesses a higher degree of variation. The clock model was selected according to the marginal likelihood estimated (MLE) from stepping stone and path sampling (Baele et al. 2012) in BEAST, with 100 path steps, 1000,000 iterations, and sampling every 2000 generations. A GTR + I + G model of base pair substitution and an uncorrelated relaxed log-normal molecular clock with a prior coalescent tree of constant size were used. A normal distribution was employed, and default values were used for the other priors. A well-accepted *Cytb* evolutionary rate was also included (Oshida et al. 2005; Nabholz et al. 2008; Ben Faleh et al. 2012a; Suzuki et al. 2015). This rate was drawn from a normal distribution with a fixed mean of 2% per million years (Ma) and a fixed standard deviation of 0.005, as employed in previous studies on jerboas (Ben Faleh et al. 2012a, b). The MCMC analysis was run twice for 20 million generations, with sampling every 1000 generations. Effective sample size values (ESS > 200) were obtained using Tracer 1.6 (Rambaut et al. 2014). TREEANNOTATOR 1.8.2 in the BEAST package was employed to summarize the tree data, with the first 25% being discarded as burn-in. The tree and divergence times were displayed and edited in FIGTREE 1.4.2.

Intraspecific Delimitation and Population Structure Analyses

To estimate the species tree and test the clade divergence revealed in the Bayesian tree, concatenated mtDNA and nDNA sequences were analyzed using *BEAST (Heled and Drummond 2010), which was implemented in BEAST 1.8.2. To minimize the size of the parameter space, we enforced the topology obtained from MrBayes in all dating analyses. The substitution models were unlinked, and substitution parameters were set according to the JMODELTEST results. We chose Yule Process species tree priors in a piecewise linear and constant root population size model. Nuclear gene clock models were assessed using the *ucld.stdev* parameter based on preliminary runs in which the uncorrelated relaxed log-normal clock prior was applied and the resulting distribution examined, as recommended in the program documentation (Drummond et al. 2007; Kohli et al. 2015). The following suitable substitution models were selected using jModeltest: GTR + I + G (mtDNA), HKY + G (GHR), HKY + G (IRBP), and GTR + I (RAG1). The uncorrelated relaxed log-normal clock was set for all loci, and the mutation rates were set to 2% (mtDNA), 0.17% [*GHR* (Galewski et al. 2006)], 0.083% [*IRBP* (Nabholz et al. 2008)], and 0.29% [*RAG1* (Townsend 2007)]. The MCMC chains were run for 200 million generations, with sampling every 5000 generations. The convergence of the MCMC chains was examined in Tracer 1.6, and the first 25% were discarded as burn-in. The trees and posteriors were displayed and edited in FIGTREE 1.4.2.

To test the species delimitation and identify the population structure, we performed three analyses. First, we employed BAPS 6 (Corander et al. 2008) to analyze the genetic structure by clustering sampled individuals into groups for both the mtDNA and nDNA sequences. We conducted clustering with analysis of linked loci in the population mixture analysis (Corander and Tang 2007), and other parameters were set to the default values. Then, ten replicates of K (from two to 15) were generated to determine the optimal number of groups, which was identified based on the highest marginal log likelihood. Second, we identified the population structure in three nuclear loci in STRUCTURE 2.3.4 (Pritchard et al. 2000) with a burn-in of 200,000 followed by two million MCMC iterations of an admixture model with correlated allele frequencies. The number of clusters (K) was set from 1 to 10 to cover the numbers of main clades that could be identified in the mtDNA trees (see Results), and each K was analyzed five times. The K with the highest likelihood was identified in Structure Harvester (Earl and Vonholdt 2012). The merging of the replicated runs for K was implemented in CLUMPP (Jakobsson and Rosenberg 2007), and the results in bar plots were visualized in DISTRUCT (Rosenberg 2004). Third, we used BPP 3.2 (Yang and Rannala 2014) and ran reversible-jump MCMC (rjMCMC) analyses for 500,000 generations (with a sampling interval of five), with a burn-in phase of 100,000. The priors were set to $\theta \sim G$ (20, 2, 000) and $\tau \sim G$ (20, 1, 000). The other divergence time parameters were assigned to the Dirichlet prior (Yang and Rannala 2014). Each analysis was run twice with different starting seeds to ensure consistency.

The population structure based on concatenated mtDNA sequences was tested in Arlequin v3.5.2.2 (Excoffier and Lischer 2010) by analysis of molecular variance (AMOVA).

We tested different grouping arrangements based on the phylogenetic results. The statistical significance of the variance components in the AMOVA was tested with 10,000 permutations. To determine the relationship between the genetic variation patterns among the populations and the natural geographical barriers, we employed BARRIER 2.2 (Manni et al. 2004) to identify the potential genetic barriers based on the maximum difference algorithm (Monmonier 1973). To decrease the deviation caused by the number of individuals, we selected sample sites with more than three individuals. Pairwise F_{ST} was calculated using the corrected average pairwise difference [PiXY-(PiX + PiY)/2] in Arlequin v3.5.2.2. The geographical coordinates of each population were connected by Delaunay triangulation, and the corresponding Voronoi tessellation was derived.

Demographic Histories

For the demographic history, we utilized three common methods. First, Tajima's *D* (Tajima 1989) and Fu's *Fs* (Fu 1997) were conducted to test the neutrality and demographic history for the six clades based on combined mtDNA genes in Arlequin v3.5.2.2, with 10,000 bootstrap replicates. Significant negative values for Tajima's *D* and Fu's *Fs* generally indicate a recent demographic expansion. Second, mismatch distributions of pairwise sequences were simulated for the clades using Arlequin, with 10,000 bootstrap replicates. The raggedness index (Rag), representing smoothness, and the sum of squares deviation (SSD), representing the degree of fit, were calculated to determine whether the sequences deviated significantly from a model of population expansion. Unimodal mismatch distributions and small Rag values with non-significant *p*-values for the SSD imply recent demographic expansion or range expansion for a population (Li et al. 2015), whereas stable populations are characterized by ragged and erratic mismatch distributions (Rogers and Harpending 1992). Third, extended Bayesian skyline plots (EBSP) were implemented in BEAST for combined mtDNA sequences to estimate effective population size changes on an evolutionary time scale. For the parameters other than the prior tree, the settings were same as in the divergence time analysis. The prior tree was set as the coalescent EBSP and sampled every 1000 steps for all 20 million steps. The MCMC convergence and ESS values for all the parameters were assessed in Tracer v1.6.

Data Availability Statement

All datasets generated and/or analyzed in the current study are available from the corresponding author upon reasonable request.

Results

Phylogenetic Relationships of *Dipus sagitta*

For mtDNA sequencing, 1113 bp *Cytb* and 661 bp *COI* sequences were obtained from 203 and 204 individuals, respectively. For the nDNA sequences, *GHR* (631 bp), *IRBP* (1186 bp), and *RAG1* (826 bp) were obtained from all 204 individuals. The sequence dataset generated herein is available in the GenBank repository with Accession numbers KX399483-KX399764. For the entire population, 158 polymorphic sites and 115 haplotypes were identified in the mtDNA sequences. The obtained values of haplotype diversity (Hd) and nucleotide diversity [Pi (%)] were 0.973 and 5.489%, respectively (Table 1). In the six clades, the Hd values ranged from 0.675 to 0.996, whereas the Pi values ranged from 0.256 to 1.104%. The DNA polymorphism indices for each gene are listed in Supplementary Table 3.

Phylogenetic analyses of the mtDNA sequences reconstructed a well-resolved Bayesian tree with six major clades (labelled A–F; Fig. 2a), which were strongly associated with geographic areas. Clade A encompassed the broadest range and was distributed in the deserts adjoining the Hexi Corridor and the Qilian Mountains; Clade B was confined to the area to the east of the Helan Mountains and to the south of the Yinshan Mountains; Clade C included individuals to the north of the Yinshan Mountains; Clade D corresponded to populations living in sandy lands on both sides of the Daxing'anling Mountains; Clade E included individuals from the Gurbantunggut Desert in the Junggar Basin isolated by the Tianshan and Altay mountains; and Clade F was restricted to the Tarim Basin by the Tianshan Mountains to the north and the QTP to the south (Fig. 1). However, the three nDNA genes lacked structure (except in Clade F), due to a slow substitution rate of nuclear genes (Supplementary Fig. 1).

The phylogenetic structure was also supported by the MJ networks, which generated six groups of haplotypes as well, although these groups did not fully correspond to the phylogenetic tree (Fig. 2b). In addition, 115 haplotypes of the combined mtDNA sequences were identified, and shared haplotypes were found among Clades A, B, and C. Certain haplotypes from different clades also co-occurred in contact zones. The haplotype group corresponding to Clade A showed a star-like structure centered on the haplotype that was the most frequent haplotype of the MJ network. The haplotype groups corresponding to Clade B, C, and F also exhibited less apparent star-like shapes (Fig. 2b). The mean percentage of genetic distance (K2P) between the six clades ranged from 0.0221 ± 0.0030 to 0.1275 ± 0.0085. Within-group pairwise K2P distances ranged from 0.0048 ± 0.0009 (Clade D) to 0.0201 ± 0.0015 (Clade E) (Supplementary Table 4).

Table 1 Genetic diversity estimates based on mitochondrial DNA for all individuals and each clade

All/Clade	n	S	h	Hd	Pi (%)	Tajima's *D*	Fu's *Fs*	Hr	SSD
All	237	158	115	0.973	5.489	1.025	0.189	0.003	0.009
A	120	175	75	0.993	0.625	**−2.289**	**−24.357**	0.002	0.003
B	20	102	19	0.995	1.104	−1.366	**−4.240**	0.009	0.021
C	24	83	20	0.975	1.013	−0.779	−3.413	0.009	0.009
D	18	10	6	0.675	0.256	−0.983	26.925	**0.204**	**0.397**
E	12	33	10	0.955	0.500	−0.854	−1.775	0.022	0.033
F	43	96	39	0.996	0.715	**−1.614**	**−21.912**	0.003	0.002

Sampled localities, number of individuals (*n*), number of polymorphic sites (*S*), number of haplotypes (*h*), haplotype diversity (*Hd*), nucleotide diversity (*Pi*), Harpending's raggedness index (*Hr*) and sum of squared deviations (*SSD*). Column n includes the data (only *Cytb*) from GenBank, although these data were excluded in the genetic diversity analysis to avoid bias. Values in bold indicate significance based on the *p*-value

Divergence Time of the Six Clades

The divergence time data indicated that the common ancestor of *D. sagitta* and the outgroup species (*Jaculus jaculus*, which was the closest species in our tree) diverged approximately 13.38 Ma (Fig. 2c). A hierarchic pattern occurred in the entire in-group. Clade F first diverged from the other clades approximately 7.57 Ma [2.39–25.43 Ma, 95% highest posterior density (HPD)], placing the ancestry of all living representatives of the species in the late Miocene. Clade E appeared approximately 5.54 Ma (1.71–18.55 Ma, 95% HPD). Both Clades E and F diverged during the late Miocene to the Pliocene.

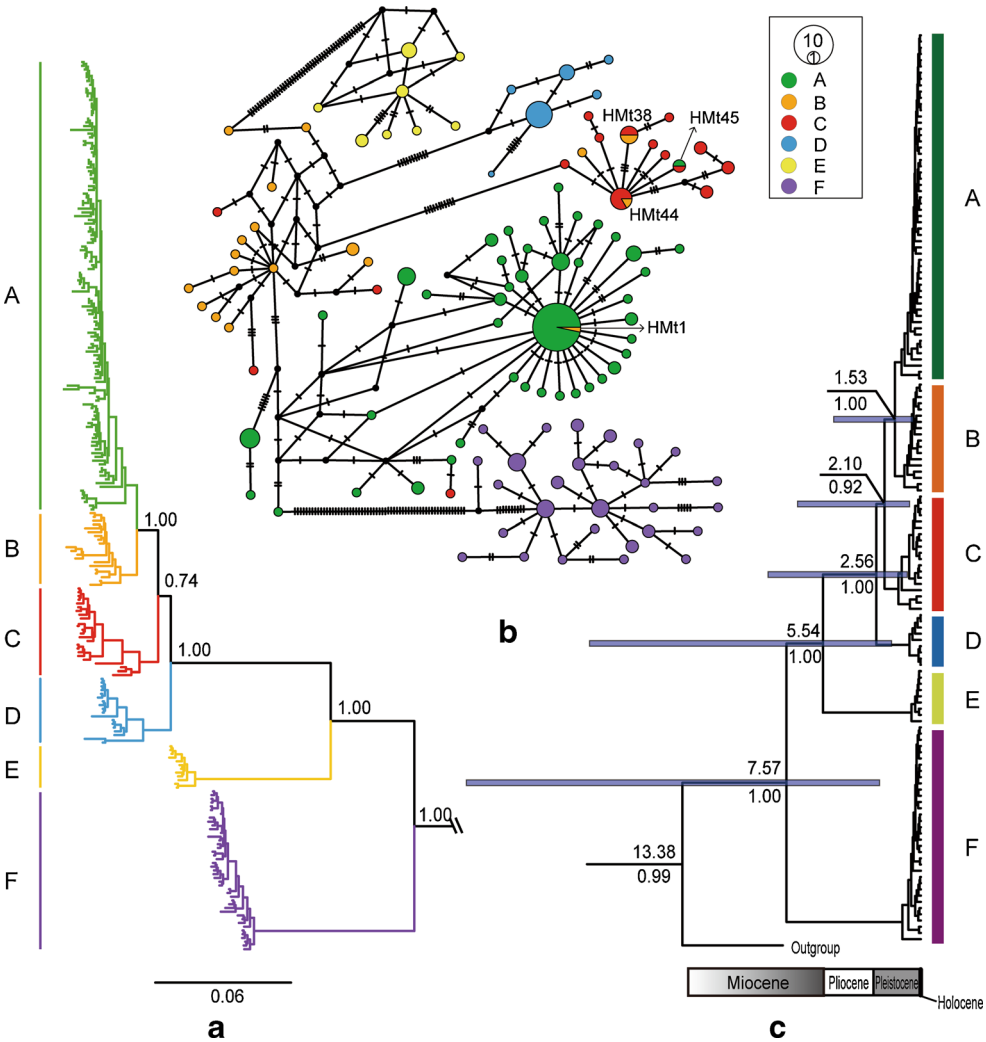

Fig. 2 Phylogenetic tree, haplotype network, and divergence time estimation based on the mtDNA haplotypes of *D. sagitta*. **a** Phylogenetic tree from the Bayesian inference analysis; the Bayesian posterior probabilities are shown below the branches. **b** Median-joining network with node sizes proportional to the frequencies of haplotypes. The number of hatch marks between haplotypes is proportional to the number of mutational changes. Dots represent undetected haplotypes. **c** Values above the branches represent the divergence time estimates and 95% highest posterior density under 0.01 substitutions per million years, and values below branches indicate Bayesian posterior probabilities

Subsequent divergence events occurred starting at the beginning of the Quaternary, with Clades D, C, and B separating approximately 2.56 Ma (0.84–8.60 Ma), 2.10 Ma (0.71–6.96 Ma), and 1.52 Ma (0.52–4.94 Ma), respectively (Fig. 2c).

Intraspecific Delimitation and Population Structure

According to the BAPS analysis, six and two genetically structured groups were optimally partitioned based on the mtDNA and nDNA data, respectively (Fig. 3a). STRUCTURE analyses of all nuclear loci identified the best clustering value at K = 2. When K = 2, the bar plots showed a consistent pattern in which Clade F formed an independent clade and the other clades showed a weak differentiation of nuclear genes (Fig. 3b). The *BEAST analyses showed a consistent intraspecific tree with all sequence data (Fig. 3c). In the BPP species delimitation analysis, six clades were strongly supported (posterior = 1.0; Fig. 3c). However, the structure of the phylogenetic tree was inconsistent with the *BEAST tree. The posterior probabilities were low for Clades A–D, whereas high support values (>95) were observed for Clades E and F.

AMOVA showed a maximum percentage of variance (80.60%) when the populations were grouped into six clades (Table 2). The variance between sampled localities within clades and within sampled localities was 5.10% and 14.30%, respectively. The analyses using BARRIER 2.2 identified eight main putative genetic barriers based on the F_{ST} matrices with mtDNA (Fig. 4). The eight putative barriers separated all sampled localities into six regions that were coincident with the phylogenetic structure. In the nDNA data, the same geographical barrier was observed between Clade F and the other clades based on F_{ST} values (Supplementary Table 5), whereas the other geographical barriers varied randomly (Fig. 4).

Demographic History

The results regarding the demographic histories of each clade were as follows. The six clades showed entirely different results for Tajima's *D* and Fu's *Fs* (Table 1). For Clades A and F,

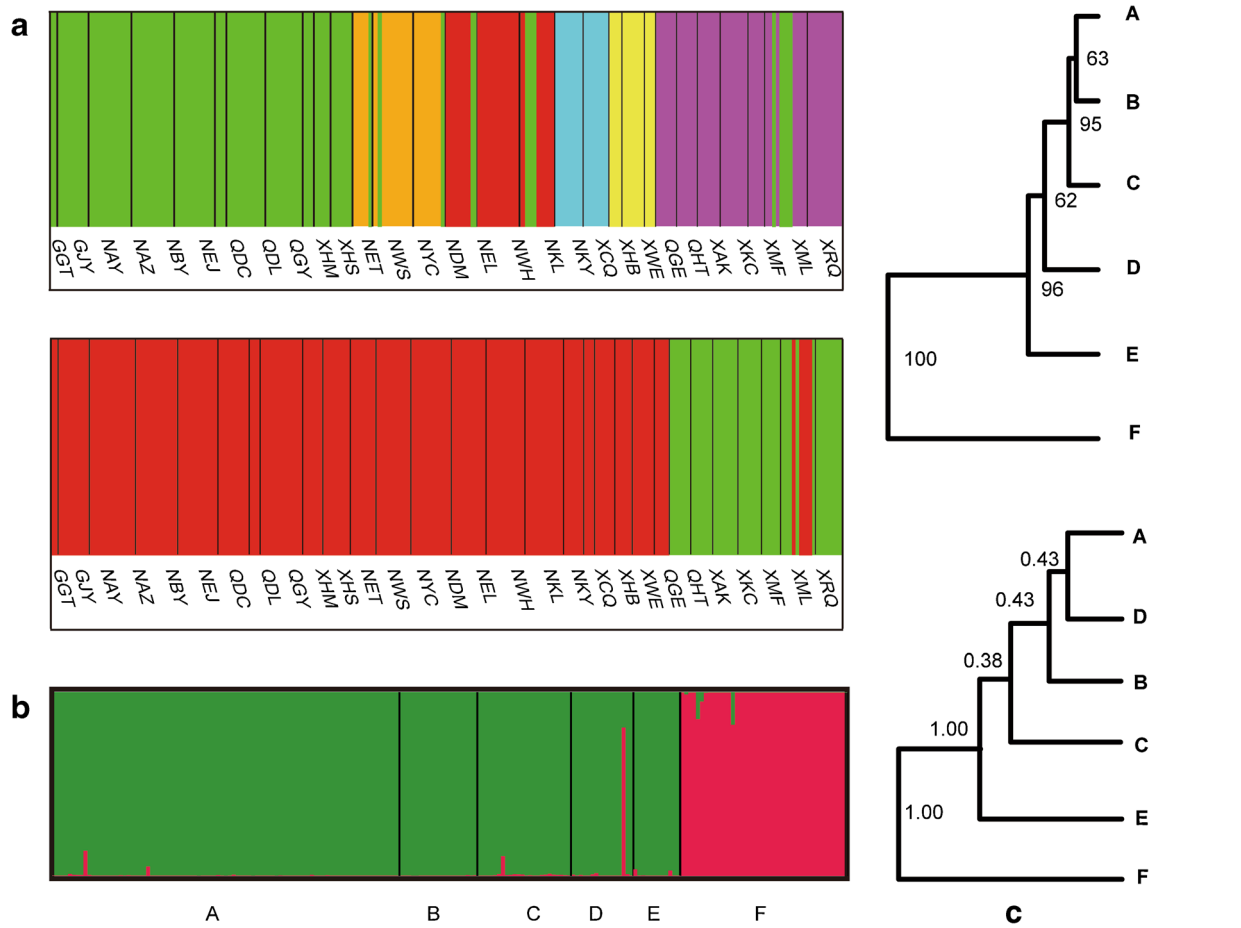

Fig. 3 Population structure and intraspecific phylogenies based on mtDNA and nDNA sequences. **a** BAPS bar plots (each column represents a sampling site separated by a white line) based on mtDNA (above) and nDNA sequences (below). **b** STRUCTURE bar plot of the best K value (K = 2) based on nuclear sequences. Codes under the figures represent the six clades. **c** Coalescent species tree estimated by *BEAST (above) and BPP (below) based on all sequences. Posterior probabilities are shown at the nodes

Table 2 Analysis of molecular variance (AMOVA) based on pairwise differences for mitochondrial DNA from *Dipus sagitta*

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation (%)	Fixation index ($p < 0.005$)
Among six clades	5	7639.538	49.471***	80.60	FCT = 0.806***
Among sampled localities within clades	23	686.383	3.129***	5.10	FSC = 0.263***
Within sampled localities	171	1501.273	8.779***	14.30	FST = 0.857***

the values of Tajima's *D* and Fu's *Fs* were negative, with significant *p* values, indicating demographic expansion. Clades B, C, D, and E mainly presented negative, but not significant, Tajima's *D* and Fu's *Fs* values, with the only exception being the Fu's *Fs* value (significantly negative) of Clade B. The mismatch distribution analysis of Clades A and F suggested a clear recent history of population expansion, as indicated by the typical smooth, unimodal distribution (Fig. 5a), consisting of low and non-significant SSD and Rag values (Table 1). Clades B, C, and E exhibited a good model fit in terms of SSD and Rag, but the values were slightly higher than those of Clades A and F. Clade D was characterized by ragged and erratic mismatch distributions, indicating a scenario of demographic equilibrium or decline. The demographic scenario for *D. sagitta* determined through EBSP analysis suggested different patterns for different clades. The EBSP analyses (Fig. 5b) showed population expansion for Clades B and C since the beginning of the middle Pleistocene, when the climate became much drier and colder than in the early Pleistocene. Clade A increased suddenly during the penultimate glaciation, and the populations of Clades E and F expanded at the same time. Clades A, B, and F increased considerably, whereas Clades C and E grew only slightly. In contrast, Clade D showed a stable condition in terms of the effective population size.

Discussion

Geographic Barriers and Climatic Demarcation Shaping Genetic Structure

The results of our mtDNA analyses, supported by the phylogenetic and MJ network analyses, strongly suggested that six major clades of *D. sagitta* occurred across its distribution area in inland East Asia. The six *D. sagitta* clades were strongly associated with the geographical terrain (see Fig. 1 and Results), and our results were generally consistent with those of previous studies on certain subspecies (Shenbrot et al. 2008). This phylogeographic pattern has previously been partially described for other vertebrate species, such as the midday gerbil (*Meriones meridianus*) (Wang et al., 2013) and the green toad (*Bufo viridis* subgroup) (Zhang et al. 2008) because of the broad distribution range we sampled.

Based on BARRIER analysis, eight boundaries were highly correlated with certain large mountain ranges and important climate division lines (Fig. 4). The combined barrier lines separating Clade F from Clade E corresponded to the Tianshan Mountains, which is an important geographical barrier interrupting gene flow in various animal taxa, such as fishes (Li et al. 2016), amphibians (Zhang et al. 2008), reptiles (Macey et al. 1999; Pang et al. 2003), and mammals (Mahmut

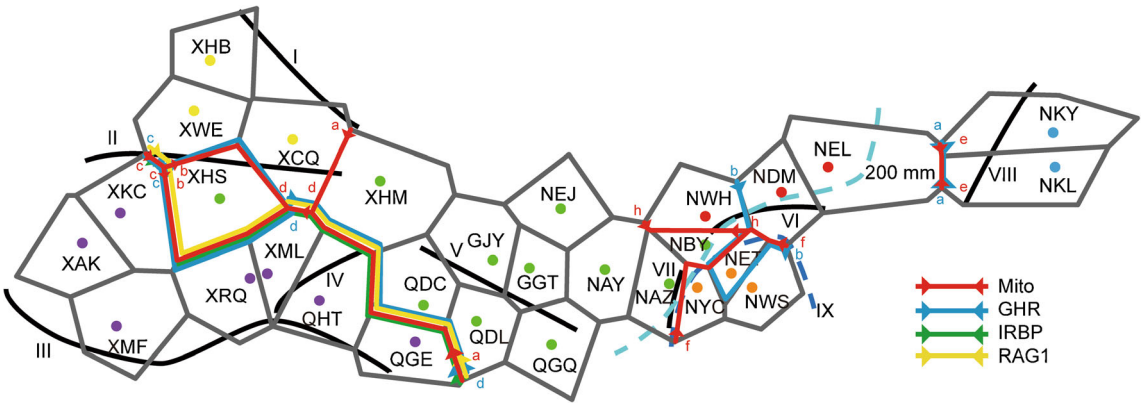

Fig. 4 Geographical barriers identified through the analysis of mitochondrial and nuclear loci of *D. sagitta*. The coloured stripe indicates the major barrier determined by the maximum difference algorithm in BARRIER 2.2 based on F_{ST} differences. Major geographical features in the study area are indicated by numbers: *I*, Altay Mountains; *II*, Tianshan Mountains; *III*, Kunlun Mountains; *IV*, Altun Mountains; *V*, Qilian Mountains; *VI*, Yinshan Mountains; *VII*, Helan Mountains; *VIII*, Daxing'anlling Mountains; and *IX*, Yellow River. The light blue dashed line represents the 200 mm precipitation line

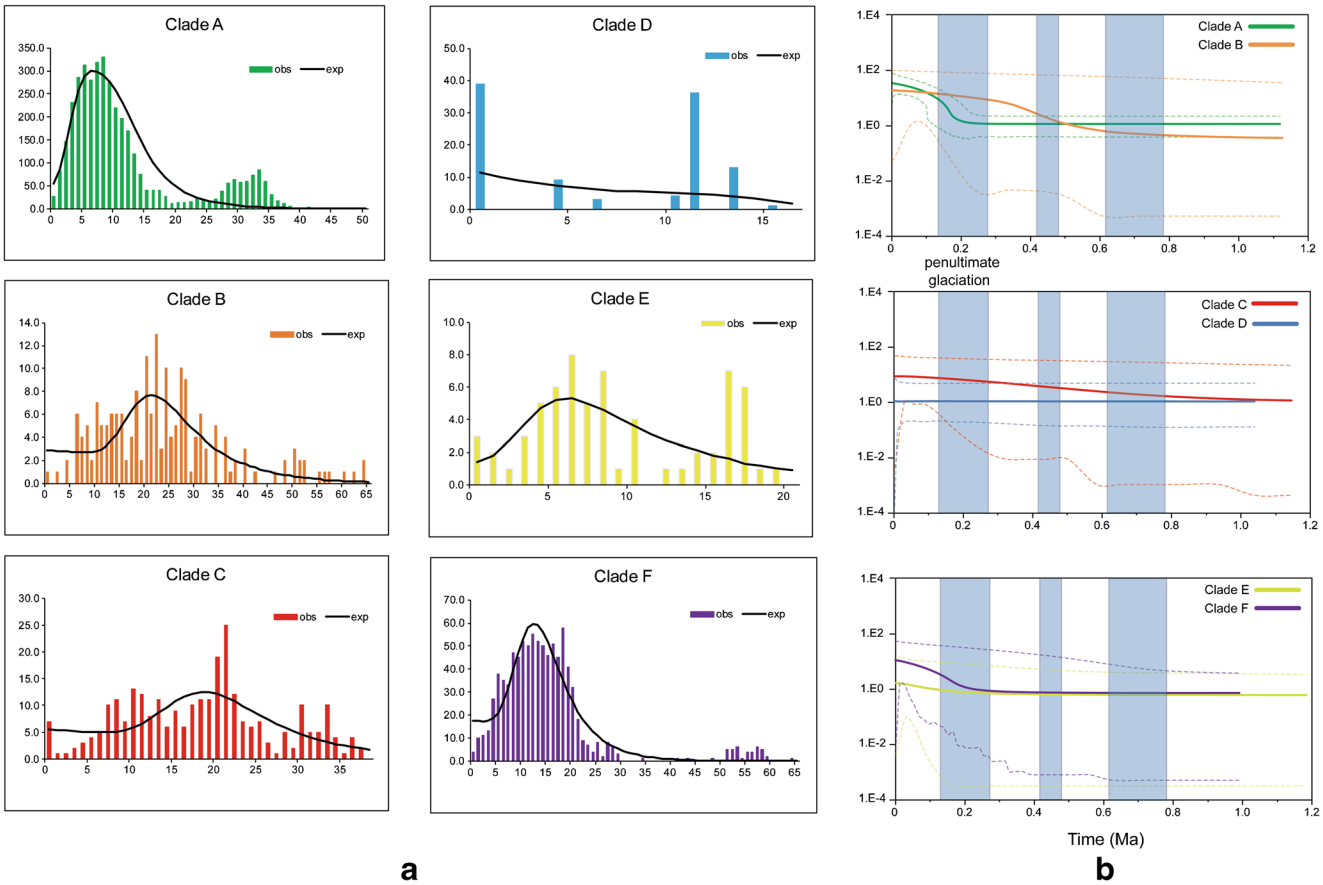

Fig. 5 Pairwise mismatch distribution (**a**) and extended Bayesian skyline plots (**b**) depicting the demographic history of each clade. For mismatch distributions, coloured bars represent the observed distribution of pairwise differences, and black lines represent the theoretical expected distribution under a population expansion model. For the extended Bayesian skyline plots, coloured solid lines indicate the median value of effective population size, whereas dotted lines represent the upper and lower 95% confidence intervals. The x-axis of the skyline plots is time in millions of years before the present, and the y-axis is the estimated effective population size (N_e)

et al. 2002; Ludt et al. 2004). Although Clades A and F exhibited a contact zone, they did not share a haplotype and presented a large number of mutational changes (Fig. 2b). Clades A and F apparently could not span the Qaidam Basin (Fig. 1). We inferred that paleo-lake appearing in the Pliocene (Zhang 2012) and a salt lake and saline wasteland that formed in central Qaidam in the Pleistocene obstructed the gene flow between the northern and southern sides of this basin (Wang et al. 2009). A recent study of another jerboa (*Jaculus orientalis*) in Tunisia also illustrated that Chott El Jerid has acted as a geographical barrier and induced strong genetic differentiation (Ben Faleh et al. 2016). Additionally, the phylogenetic topological structure was inconsistent among the different analyses when we combined the nuclear DNA at a slow substitution rate, with Clade F being located at stable positions in the trees (Supplementary Fig. 1). This result was partly because their divergence times were much longer than those of Clades A-D and therefore accumulated more mutation sites, partly because no effective geographical barriers, such as the Tianshan Mountains, exist among Clades A-D to form effective obstructions. In addition, *D. sagitta* was found male-biased dispersal (Shenbrot et al. 2008), leading to gene exchange in nuclear genes rather than mitochondrial genes. However, Clade F showed distinct genetic distance among all sequences in every analysis. Therefore, whether this clade represents a cryptic species must be considered in future studies. Here, we continue to consider Clade F as a subspecies of *D. sagitta*.

Individuals in Clade E also display high divergence from the remaining four clades according to the MJ network (Fig. 2b) and F_{ST} values (Supplementary Table 5). The coalescent species tree estimated by *BEAST and BPP based on all sequences supported Clade E as a monophyletic group with high posterior probabilities (Fig. 3c). The major barrier separating this clade from the others is the Tianshan Mountains (Fig. 4). However, the Tianshan Mountains and Altay Mountains form a natural corridor connecting the Junggar Basin and the MP (Zhang 2011), and there are historical records of the presence of *D. sagitta* along this corridor. To determine whether a genetic connection exists between Clades E and A, additional sampling efforts in the potential contact areas will be required.

Clades A, B, C, and D are mainly distributed on the MP and adjacent regions (Fig. 1), which lack obvious terrain barriers. The haplotype network showed shared haplotypes among Clades A, B, and C, indicating that frequent gene flow occurred among the populations during their expansion (Figs. 1 and 2b). The nuclear gene analyses of these four clades also showed low posterior probabilities and unstable genetic structure (Figs. 3 and 4). However, the combined feature formed by the western boundary of the Hunshandake Sandy Land, the Yinshan Mountains, and the Helan Mountains is coincident with the 200 mm precipitation line. This isohyet has been observed to act as a dividing line in several taxa, such as *Nanorana parkeri* (Zhou et al., 2014) and *Hippophae tibetana* (Wang et al., 2010). Compared with Clades E and F, which inhabit the driest regions with annual precipitation values of less than 100 mm, Clade A covers the widest ecological extent with regard to humidity and elevation, and Clades B and D primarily inhabit more humid sandy lands (Fig. 1). Therefore, the ecological divergence between Clades C and D most likely induced their genetic differentiation.

In addition, most of the region of Clade B overlaps with that of Clade A, which may have been caused by individuals from Clade A introgressing into this area. However, migration in the opposite direction did not occur because of the barrier of the Helan Mountains (Fig. 1). Moreover, the ecological breadth of Clade A may have been wider than that of Clade B, which inhabits wetter areas. Therefore, the Helan Mountains have acted as an important climate boundary and prevented the dispersal of Clade B. *Meriones unguiculatus* shows a similar pattern in the same area (Liang et al. 2007). However, to clarify how the phylogeographic clades have adapted to local environmental conditions, further investigations must be performed to analyze the genetic evidence from related functional genes and the correlations between ecological divergence and phylogenetic splits.

Divergence Time and Biogeographic Scenario

Arid zone has continuously occupied most of inland East Asia since the Oligocene. However, the causes of aridification shifted from a subtropical high planetary wind system to an inland continental arid system at the boundary of Oligocene and Miocene. Since the early Miocene period, the extent of aridification in this area has increased due to the uplift of a series of mountains, which changed the natural geographic environment (Zhang and Guo 2005). A recent study showed that extant Dipodinae species adapted to open and arid habitats responded to these changes by expanding their distribution area from Central Asia to inland East Asia (such as Mongolia and the Gobi-Taklimakan deserts) (Pisano et al. 2015). The estimated divergence time (13.38 Ma) of the common ancestor of *D. sagitta* and *J. jaculus* is close to a previous estimate (12.65 Ma) by Pisano et al. (2015). The first fossils of the genus *Dipus* were found in Central Inner Mongolia in the middle of the late Miocene (7.50–9.10 Ma) (Qiu et al. 2006). The expansion of desert-steppe environments from the late Miocene to the early Pliocene in inland Eurasia and more humid conditions in eastern Asia offered abundant suitable habitats for the dispersal of the ancestors of *D. sagitta* (Fortelius et al., 2006; Deng et al., 2015). According to the divergence time estimate (Fig. 2c), this process can be divided into two stages: orogenic events induced early aridification and geographical isolation, and alternate environments induced ecological divergence in subsequent diversification.

In the dispersal process of the ancestor of extant *D. sagitta*, Clade F appeared at approximately 7.57 Ma. The Tarim Basin formed a suitable habitat for *D. sagitta* starting at 5.3 Ma because of the uplift of the Kunlun Mountains (Sun et al. 2008), which subsequently promoted the dispersal of Clade F. The Tianshan Mountains were continuously uplifted from the late Miocene to the early Pleistocene (7–2.58 Ma), reaching elevations similar to the current elevation (Sun et al. 2004), and these mountains interrupted genetic communication between Clades F and E. The Altay Mountains were subsequently uplifted, but their elevations were not as great as those of the Tianshan Mountains (Guo et al. 2006). Nevertheless, these mountains may have acted as barriers to Clade E in the northeast. Although the modern Gurbantunggut Desert began to form at 0.8 Ma (Shi et al., 2006a), the aridification in the Junggar Basin began in the Pliocene (Wang 1997) or even in the late Miocene, and sporadic desert areas formed and offered habitat for Clade E. The orogeny produced a significant isolating effect on *D. sagitta*, indicating that the early clade divergence was generally related to the uplift of huge mountains during the geologic history of the area. We inferred that the fundamental reason for the observed isolation was related to a barrier habitat in a high-elevation area with more precipitation, which generated forest vegetation types.

The southern MP is located at some distance from the QTP and was less influenced by the early uplift of the QTP. The process of steppification accompanied by climatic aridification in the southern MP began at the end of Oligocene and persisted until the Pliocene (Zhang 2012). Since the Qingzang Movement at 3.6–1.7 Ma (Li 1999), the modern monsoon climate has intensified and promoted a dry-cold climate on the MP. A forest-steppe environment combined with a lake-basin system transformed into typical steppe, desert steppe, and desert systems (Zhang 2012). The large-scale expansion of suitable habitats along the southern MP and the lack of high mountains formed an area with abundant opportunities for the dispersal of *D. sagitta*. The divergence among Clades A, B, C, and D began at c. 2.56 Ma (Fig. 2c), coinciding with phase B (2.6 Ma) of the Qingzang movement, which intensified the effect of the winter monsoon and led to prominent aridification on the MP (Li et al. 2002). During the Kunlun-Yellow River

Movement at 1.2–0.6 Ma (Li 1999), the Badain Jaran Desert, Tengger Desert, and Mu Us Sandy Land appeared contemporaneously at 1.2–0.8 Ma (Wang et al. 2015), offering opportunities for expansion. Compared with Clades E and F, the remaining four clades experienced more rapid diversification within 1.03 Myr (2.56–1.53 Ma) (Fig. 2c). During this period, repeated environmental changes might have generated fragmented suitable habitats, especially in the areas of Clades C and D, which induced genetic differentiations. This genetic differentiation was further reinforced and accumulated because of ecological divergence over time (Qu et al. 2015). This finding indirectly supports the existence of weak geographical isolation during later diversification and a change in suitable habitat induced by climatic changes. The divergence between Clades B and A and the overlap of the two clades may have been caused by repeated connections and interruptions of a paleo-lake located in the midstream section of the Yellow River during the formation of the ancient Yellow River (1.6–0.15 Ma) (Li et al. 2005). When the paleo-lake contracted, individuals from Clade A introgressed into this area, and gene flow was restored.

In addition, AMOVA showed that the extent of the divergence between clades (80.60%) was considerably greater than that within each clade (5.10%) or within the sampled localities (14.30%), indicating that the extant genotypes appeared rapidly over a relatively short period of time (Table 2 and Fig. 2c). Thus, early regional aridification likely promoted dispersal and clade differentiation, whereas late desert formation induced the expansion of suitable habitats and demographic structure.

Demographic History

The genetic diversity and demographic analyses presented different demographic scenarios in some clades. High haplotype diversity values and low nucleotide diversity values always suggest rapid population growth with recent differentiation (Avise 2000; Rivera et al. 2016). In the present study, the genetic diversity of Clades A and F showed this pattern (Table 1). The mismatch distribution was unimodal (Fig. 5a), and the star-shaped network and significant negative values for both Tajima's *D* and Fu's *Fs* tests also supported demographic expansion (Table 1). A similar pattern of nucleotide and haplotype diversity and demographic history has been reported for populations of *M. meridianus* in the same geographic area (Wang et al. 2013). The EBSP analysis corroborated the occurrence of population expansion, showing that the N_e of Clades A and F increased suddenly during the penultimate glaciation (Fig. 5b). Although Clade E showed high haplotype diversity and low nucleotide diversity, the remaining demographic analyses still indicated weak population expansion (Table 1 and Fig. 5a). However, the increase in time based on the EBSP analysis of Clade E was coincident with that of Clade F. Clades B and C featured similar demographic patterns. High values of Hd and Pi supported stable populations with a large long-term N_e, whereas the mismatch distribution and EBSP analyses indicated moderate population growth (Table 1 and Fig. 5b). The significant negative values of Fu's *Fs* test obtained for Clade B supported demographic expansion. In the EBSP results, the N_e of Clades B and C increased initially at the beginning of the middle Pleistocene, which was earlier than the other clades. Clade D apparently experienced a genetic bottleneck involving low haplotype diversity and nucleotide diversity, ragged and erratic mismatch distributions, and significant SSD and raggedness values, but a stable N_e according to EBSP analysis.

Although most parts of northern China were not covered by an extensive ice sheet (Shi et al., 2006a, b; Liu et al. 2016), climatic oscillations coupled with repeated expansions and contractions of desert areas likely affected geographical patterns of genetic diversity and demographic history (Hewitt 2000; Meng et al. 2015). As psammophilic organisms, *D. sagitta* are thought to have experienced interglacial contraction and glacial expansion in response to cold-warm climatic cycles. Our hypothesis was partly confirmed in the EBSP analysis. Five clades of *D. sagitta* showed demographic expansion to different degrees in the glacial period in the middle-late Pleistocene. However, the continuous expansion during the climate oscillation period appeared to be more closely related to the increasing aridification caused by orogenic movement than climate oscillation, which was in agreement with evidence from other desert species, such as *M. meridianus* (Wang et al., 2013) and *Phodopus roborovskii* (Lv et al., 2016). In addition, the easternmost part of inland East Asia (occupied by Clade D) has been sensitive to environmental changes and underwent a habitat transition from sandy land to forest during the warm period in the late Pleistocene (Zhang 2012).

Conclusion

The present study suggests that the evolutionary history of *D. sagitta* has witnessed aridification since the Late Miocene and the subsequent desertification of inland East Asia caused by the uplift of the QTP since the Pliocene. Orogenic and environmental changes separated the inland East Asia region into several distinguishable areas, and the ancestors of *D. sagitta* were affected by regional aridification as well as these geological events; thus, they diverged into different clades. The historical dynamics of desertification coincided with the genetic divergence and phylogeographic patterns of *D. sagitta*. Great mountains (such as the Tianshan Mountains) and important climate dividing lines (such as the 200 mm isohyet) were inferred to represent genetic barriers among the six clades. The orogenic history showed significant

isolating effects, and early clade divergence was generally related to the uplift of high mountains in the geologic history. During subsequent diversification, repeated environmental changes might have generated suitably fragmented habitats to induce genetic differentiation, which was then further reinforced and accentuated by ecological divergence over time. The demographic analyses suggested different demographic histories in distinct clades; however, the continuous expansion during the climate oscillation period appeared to be more closely related to the increasing aridification caused by orogenic movement than to climate oscillation.

Acknowledgements We thank the editors and anonymous reviewers for providing valuable comments. We sincerely thank Yuanbao Du and Yongbin Chang for their help with both the fieldwork and laboratory work. We thank Gregory I. Shenbrot for his suggestions and prior work on jerboas. We also thank the researchers who submitted sequences from previous studies to GenBank. This research was supported by grants from the National Natural Science Foundation of China (31172065 to L.X.), Newton Advanced Fellowship of the Royal Society of the United Kingdom (Ref. NA150142 to D.G.), and the Key Laboratory of Zoological Systematics and Evolution of the Chinese Academy of Sciences (No. Y229YX5105).

Author Contributions Q.Y., L.X. and J.C. conceived of the study; Q.Z., L.X., and L.L. collected samples; J.C. performed the laboratory work; J.C., X.L. and D.G. analysed the data; and J.C. and X.L. wrote the first draft of the manuscript. All authors contributed to the interpretation of the results and commented on the final version of this manuscript.

Compliance with Ethical Standards

Competing Interests The authors declare that they have no competing interests.

References

Avise JC (2000) Phylogeography: The history and formation of species. Harvard University Press, Cambridge

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol Biol Evol 29(9): 2157–2167. doi:10.1093/molbev/mss084

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16(1): 37–48

Batsaikhan N, Avirmed D, Tinnin D, Tsytsulina K (2016) *Dipus sagitta*. The IUCN red list of threatened species 2016: E.T6705A22202817

Ben Faleh A, Granjon L, Tatard C, BoratyŃSki Z, Cosson JF, Said K (2012a) Phylogeography of two cryptic species of African desert jerboas (Dipodidae: Jaculus). Biol J Linn Soc 107(1): 27–38. doi:10.1111/j.1095-8312.2012.01920.x

Ben Faleh A, Granjon L, Tatard C, Othmen AB, Said K, Cosson JF (2012b) Phylogeography of the greater Egyptian jerboa (*Jaculus orientalis*) (Rodentia: Dipodidae) in Mediterranean North Africa. J Zool 286(3): 208–220. doi:10.1111/j.1469-7998.2011.00868.x

Brito JC, Godinho R, Martinez-Freiria F, Pleguezuelos JM, Rebelo H, Santos X, Vale CG, Velo-Anton G, Boratynski Z, Carvalho SB, Ferreira S, Goncalves DV, Silva TL, Tarroso P, Campos JC, Leite JV, Nogueira J, Alvares F, Sillero N, Sow AS, Fahd S, Crochet PA, Carranza S (2014) Unravelling biodiversity, evolution and threats to conservation in the Sahara-Sahel. Biol Rev Camb Philos Soc 89(1): 215–231. doi:10.1111/brv.12049

Castoe TA, Spencer CL, Parkinson CL (2007) Phylogeographic structure and historical demography of the western diamondback rattlesnake (*Crotalus atrox*): A perspective on north American desert biogeography. Mol Phylogenet Evol 42(1): 193–212. doi:10.1016/j.ympev.2006.07.002

Chen KM, Abbott RJ, Milne RI, Tian XM, Liu JQ (2008) Phylogeography of *Pinus tabulaeformis* Carr. (Pinaceae), a dominant species of coniferous forest in northern China. Mol Ecol 17(19): 4276–4288. doi:10.1111/j.1365-294X.2008.03911.x

Corander J, Tang J (2007) Bayesian analysis of population structure based on linked molecular information. Math Biosci 205(1): 19–31. doi:10.1016/j.mbs.2006.09.015

Corander J, Marttinen P, Siren J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9: 539. doi:10.1186/1471-2105-9-539

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: More models, new heuristics and parallel computing. Nat Methods 9(8): 772. doi:10.1038/nmeth.2109

Deng T, Wang XM, Wang SQ, Li Q, Hou SK (2015) Evolution of the Chinese Neogene mammalian faunas and its relationship to uplift of the Tibetan Plateau. Advances in Earth Science 30(04): 407-415. doi:10.11867/j.issn.1001-8166.2015.04.0407

Drummond AJ, Ho SY, Rawlence N, Rambaut A (2007) A rough guide to BEAST 1.4

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29(8): 1969–1973. doi:10.1093/molbev/mss075

Durant SM, Pettorelli N, Bashir S, Woodroffe R, Wacher T, De Ornellas P, Ransom C, Abaigar T, Abdelgadir M, El Alqamy H, Beddiaf M, Belbachir F, Belbachir-Bazi A, Berbash AA, Beudels-Jamar R, Boitani L, Breitenmoser C, Cano M, Chardonnet P, Collen B, Cornforth WA, Cuzin F, Gerngross P, Haddane B, Hadjeloum M, Jacobson A, Jebali A, Lamarque F, Mallon D, Minkowski K, Monfort S, Ndoassal B, Newby J, Ngakoutou BE, Niagate B, Purchase G, Samaila S, Samna AK, Sillero-Zubiri C, Soultan AE, Stanley Price MR, Baillie JME (2012) Forgotten biodiversity in desert ecosystems. Science 336(6087): 1379–1380. doi:10.1126/science.336.6087.1379

Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2): 359–361. doi:10.1007/s12686-011-9548-7

Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5): 1792–1797. doi:10.1093/nar/gkh340

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(3): 564–567. doi:10.1111/j.1755-0998.2010.02847.x

Faleh AB, Allaya H, Shahin AAAB (2016) Geographic patterns of genetic variation in the greater Egyptian jerboa *Jaculus orientalis* (Dipodidae, Rodentia) from Tunisia. Biochem Syst Ecol 68: 15–22. doi:10.1016/j.bse.2016.06.015

Fortelius M, Eronen J, Liu LP, Pushkina D, Tesakov A, Vislobokova I, Zhang ZQ (2006) Late Miocene and Pliocene large land mammals and climatic changes in Eurasia. Palaeogeogr Palaeoclimatol Palaeoecol 238(1–4): 219–227. doi:10.1016/j.palaeo.2006.03.042

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147(2): 915–925

Galewski T, Tilak MK, Sanchez S, Chevret P, Paradis E, Douzery EJ (2006) The evolutionary radiation of Arvicolinae rodents (voles

and lemmings): Relative contribution of nuclear and mitochondrial DNA phylogenies. BMC Evol Biol 6: 80. doi:10.1186/1471-2148-6-80

Ge XJ, Yu Y, Yuan YM, Huang HW, Yan C (2005) Genetic diversity and geographic differentiation in endangered Ammopiptanthus (Leguminosae) populations in desert regions of northwest China as revealed by ISSR analysis. Ann Bot 95(5): 843–851. doi:10.1093/aob/mci089

Guo SJ, Zhang ZC, Wu CD, Fang SH, Zhang R (2006) The Mesozoic and Cenozoic exhumation history of Tianshan and comparative studies to the Junggar and Altai Mountains. Acta Geol Sinica 80(01): 1–15

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucl Acid S 41: 95–98

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Mol Biol Evol 27(3): 570–580. doi:10.1093/molbev/msp274

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405(6789): 907–913. doi:10.1038/35016000

Jakobsson M, Rosenberg NA (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23(14): 1801–1806. doi:10.1093/bioinformatics/btm233

Kohli BA, Fedorov VB, Waltari E, Cook JA (2015) Phylogeography of a Holarctic rodent (*Myodes rutilus*): Testing high-latitude biogeographical hypotheses and the dynamics of range shifts. J Biogeogr 42(2): 377–389. doi:10.1111/jbi.12433

Lei FM, Qu YH, Song G (2014) Species diversification and phylogeographical patterns of birds in response to the uplift of the Qinghai-Tibet Plateau and Quaternary glaciations. Curr Zool 60(2): 149–161

Leigh JW, Bryant D (2015) POPART: Full-feature software for haplotype network construction. Methods Ecol Evol 6(9): 1110–1116. doi:10.1111/2041-210X.12410

Li JJ (1999) Studies on the geomorphological evolution of the Qinghai-Xizang (Tibetan) Plateau and Asian monsoon. Marine Geol Quaternary Geol 19(01): 7–17

Li ZP, Zhang WJ, Wang M, Nie HG, Yue LP (2002) Geological views on the sandy desertification of the east part of northern China. Northw Geol 35(03): 7–22

Li RQ, Qiu WL, Zhang YL, Zhang BY (2005) New understanding of the Loess Plateau. Journal of Beijing Normal University (Natural Science) 41(04): 431–436

Li J, Zhao M, Wei SC, Luo ZH, Wu H (2015) Geologic events coupled with Pleistocene climatic oscillations drove genetic variation of Omei treefrog (*Rhacophorus omeimontis*) in southern China. BMC Evol Biol 15: 289. doi:10.1186/s12862-015-0572-1

Li GG, Peng ZG, Zhang RY, Tang YT, Tong C, Feng CG, Zhang CF, Zhao K (2016) Mito-nuclear phylogeography of the cyprinid fish *Gymnodiptychus dybowskii* in the arid Tien Shan region of Central Asia. Biol J Linn Soc 118(2): 304–314. doi:10.1111/bij.12724

Liang J, Zhou LZ, Zhao TB, Zhang BW, Ning SL (2007) Genetic variation and geographicial differentiation of cytochrome b gene of clawed jird (*Meriones unguiculatus*). Acta Theriol Sinica 27(02): 138–145

Liao JC, Jing DD, Luo GJ, Wang Y, Zhao LM, Liu NF (2016) Comparative phylogeography of *Meriones meridianus*, *Dipus sagitta*, and *Allactaga sibirica*: Potential indicators of the impact of the Qinghai-Tibetan Plateau uplift. Mammal Biol 81(1): 31–39. doi:10.1016/j.mambio.2015.05.002

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452. doi:10.1093/bioinformatics/btp187

Liu SX, Jiang N, Xue DY, Cheng R, Qu YH, Li XX, Lei FM, Han HX (2016) Evolutionary history of *Apocheima cinerarius* (Lepidoptera: Geometridae), a female flightless moth in northern China. Zool Scr 45(2): 160–174. doi:10.1111/zsc.12147

Ludt CJ, Schroeder W, Rottmann O, Kuehn R (2004) Mitochondrial DNA phylogeography of red deer (*Cervus elaphus*). Mol Phylogenet Evol 31(3): 1064–1083. doi:10.1016/j.ympev.2003.10.003

Lv X, Xia L, Ge DY, Wen ZX, Qu YH, Lu L, Yang QS (2016) Continental refugium in the Mongolian Plateau during Quaternary glacial oscillations: Phylogeography and niche modelling of the endemic desert hamster, *Phodopus roborovskii*. PLoS One 11(2): e0148182. doi:10.1371/journal.pone.0148182

Macey JR, Wang Y, Ananjeva NB, Larson A, Papenfuss TJ (1999) Vicariant patterns of fragmentation among gekkonid lizards of the genus *Teratoscincus* produced by the Indian collision: A molecular phylogenetic perspective and an area cladogram for Central Asia. Mol Phylogenet Evol 12(3): 320–332. doi:10.1006/mpev.1999.0641

Mahmut H, Masuda R, Onuma M, Takahashi M, Nagata J, Suzuki M, Ohtaishi N (2002) Molecular phylogeography of the red deer (*Cervus elaphus*) populations in Xinjiang of China: Comparison with other Asian, European, and north American populations. Zoolog Sci 19(4): 485–495. doi:10.2108/zsj.19.485

Maldonado-Sanchez D, Gutierrez-Rodriguez C, Ornelas JF (2016) Genetic divergence in the common bush-tanager *Chlorospingus ophthalmicus* (Aves: Emberizidae) throughout Mexican cloud forests: The role of geography, ecology and Pleistocene climatic fluctuations. Mol Phylogenet Evol 99: 76–88. doi:10.1016/j.ympev.2016.03.014

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. Hum Biol 76(2): 173–190. doi:10.1353/hub.2004.0034

Meng HH, Gao XY, Huang JF, Zhang ML (2015) Plant phylogeography in arid Northwest China: Retrospectives and perspectives. J Syst Evol 53(1): 33–46. doi:10.1111/jse.12088

Miao YF, Herrmann M, Wu FL, Yan XL, Yang SL (2012) What controlled mid-late Miocene long-term aridification in Central Asia? - global cooling or Tibetan Plateau uplift: A review. Earth-Sci Rev 112(3–4): 155–172. doi:10.1016/j.earscirev.2012.02.003

Monmonier MS (1973) Maximum-difference barriers: An alternative numerical regionalization method. Geogr Anal 5(3): 245–261

Murphy NP, Breed MF, Guzik MT, Cooper SJB, Austin AD (2012) Trapped in desert springs: Phylogeography of Australian desert spring snails. J Biogeogr 39(9): 1573–1582. doi:10.1111/j.1365-2699.2012.02725.x

Nabholz B, Glemin S, Galtier N (2008) Strong variations of mitochondrial mutation rate across mammals–the longevity hypothesis. Mol Biol Evol 25(1): 120–130. doi:10.1093/molbev/msm248

Oshida T, Abramov A, Yanagawa H, Masuda R (2005) Phylogeography of the Russian flying squirrel (*Pteromys volans*): Implication of refugia theory in arboreal small mammal of Eurasia. Mol Ecol 14(4): 1191–1196. doi:10.1111/j.1365-294X.2005.02475.x

Pang JF, Wang YZ, Zhong Y, Hoelzel AR, Papenfuss TJ, Zeng XM, Ananjeva NB, Zhang YP (2003) A phylogeny of Chinese species in the genus *Phrynocephalus* (Agamidae) inferred from mitochondrial DNA sequences. Mol Phylogenet Evol 27(3): 398–409. doi:10.1016/S1055-7903(03)00019-8

Phillipsen IC, Metcalf AE (2009) Phylogeography of a stream-dwelling frog (*Pseudacris cadaverina*) in southern California. Mol Phylogenet Evol 53(1): 152–170. doi:10.1016/j.ympev.2009.05.021

Pisano J, Condamine FL, Lebedev V, Bannikova A, Quere JP, Shenbrot GI, Pages M, Michaux JR (2015) Out of Himalaya: The impact of past Asian environmental changes on the evolutionary and biogeographical history of Dipodoidea (Rodentia). J Biogeogr 42(5): 856–870. doi:10.1111/jbi.12476

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2): 945–959

Qiu ZD, Wang XM, Li Q (2006) Faunal succession and biochronology of the Miocene through Pliocene in Nei Mongol (Inner Mongolia). Vertebr Palasiatic 44(2): 164–181

Qu YH, Song G, Gao B, Quan Q, Ericson PGP, Lei FM (2015) The influence of geological events on the endemism of east Asian birds studied through comparative phylogeography. J Biogeogr 42(1): 179–192. doi:10.1111/jbi.12407

Rambaut A, Drummond AJ, Suchard M (2014) Tracer v1.6. <http://beast.bio.ed.ac.uk/Tracer>

Rivera DS, Vianna JA, Ebensperger LA, Palma RE (2016) Phylogeography and demographic history of the Andean degu, *Octodontomys gliroides* (Rodentia: Octodontidae). Zool J Linn Soc 178(2): 410–430. doi:10.1111/zoj.12412

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9(3): 552–569

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61(3): 539–542. doi:10.1093/sysbio/sys029

Rosenberg NA (2004) DISTRUCT: A program for the graphical display of population structure. Mol Ecol Notes 4(1): 137–138. doi:10.1046/j.1471-8286.2003.00566.x

Shenbrot GI (1991a) Geographical variation of the three-toed brush-footed jerboa, *Dipus sagitta* (Rodentia, Dipodidae). I, general patterns of intraspecific variability and subspecific differentiation in the western part of the species range. Zool Zh 70(5): 101–110

Shenbrot GI (1991b) Geographical variation of the three-toed brush-footed jerboa, *Dipus sagitta* (Rodentia, Dipodidae). II, Subspecific differentiation in the eastern Kazakhstan, Touva and Mongolia. Zool Zh 70(7): 91–97

Shenbrot GI, Sokolov VE, Heptner VG, Kovalskaya YM (2008) Jerboas: Mammals of Russia and adjacent regions. Science Publishers: Enfield, Jersey & Plymouth

Shi ZT, Fang XM, Song YG, An ZS, Yang SL (2006a) Loess sediments in the north slope of Tianshan Mountains and its indication of desertification since middle Pleistocene. Marine Geol Quaternary Geol 26(03): 109–114

Shi YF, Cui ZJ, Su Z (2006b) The Quaternary glaciations and environmental variations in China. Hebei Science and Technology Press, Shijiazhuang

Song S, Bao SJ, Wang Y, Bao XK, An B, Wang XL, Liu NF (2013) Population structure and demographic history of the chukar partridge *Alectoris chukar* in China. Curr Zool 59(4): 458–474

Song G, Zhang RY, DuBay SG, Qu YH, Dong L, Wang WJ, Zhang YY, Lambert DM, Lei FM (2016) East Asian allopatry and north Eurasian sympatry in long-tailed tit lineages despite similar population dynamics during the late Pleistocene. Zool Scr 45(2): 115–126. doi:10.1111/zsc.12148

Su ZH, Zhang ML (2013) Evolutionary response to Quaternary climate aridification and oscillations in north-western China revealed by chloroplast phylogeography of the desert shrub *Nitraria sphaerocarpa* (Nitrariaceae). Biol J Linn Soc 109(4): 757–770. doi:10.1111/bij.12088

Sun JM, Zhu RX, Bowler J (2004) Timing of the Tianshan Mountains uplift constrained by magnetostratigraphic analysis of molasse deposits. Earth Planet Sc Lett 219(3–4): 239–253. doi:10.1016/S0012-821x(04)00008-1

Sun JM, Zhang LY, Deng CL, Zhu RX (2008) Evidence for enhanced aridity in the Tarim Basin of China since 5.3 ma. Quaternary Sci Rev 27(9–10): 1012–1023. doi:10.1016/j.quascirev.2008.01.011

Suzuki Y, Tomozawa M, Koizumi Y, Tsuchiya K, Suzuki H (2015) Estimating the molecular evolutionary rates of mitochondrial genes referring to Quaternary ice age events with inferred population expansions and dispersals in Japanese *Apodemus*. BMC Evol Biol 15: 187. doi:10.1186/s12862-015-0463-5

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3): 585–595

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12): 2725–2729. doi:10.1093/molbev/mst197

Townsend JP (2007) Profiling phylogenetic informativeness. Syst Biol 56(2): 222–231. doi:10.1080/10635150701311362

Wang SJ (1997) Formation and evolution of the geographical environment of Junggar Basin since the late Cenozoic. Arid Land Geography 20: 9–16

Wang YG, Li YM, Chen ZY, Liu LF, An Y, Kang Q, Ren WY, Ma GZ (2009) Evolution of the sedimentary environment in the Qaidam Basin over the Quaternary period. Hydrogeol Engineering Geol (01): 128–132

Wang H, Laqiong, Sun K, Lu F, Wang YG, Song ZP, Wu QH, Chen JK, Zhang WJ (2010) Phylogeographic structure of *Hippophae tibetana* (Elaeagnaceae) highlights the highest microrefugia and the rapid uplift of the Qinghai-Tibetan Plateau. Mol Ecol 19(14): 2964–2979. doi:10.1111/j.1365-294X.2010.04729.x

Wang Y, Zhao LM, Fang FJ, Liao JC, Liu NF (2013) Intraspecific molecular phylogeny and phylogeography of the *Meriones meridianus* (Rodentia: Cricetidae) complex in northern China reflect the processes of desertification and the Tianshan Mountains uplift. Biol J Linn Soc 110(2): 362–383. doi:10.1111/bij.12123

Wang F, Sun DH, Chen FH, Bloemendal J, Guo F, Li ZJ, Zhang YB, Li BF, Wang X (2015) Formation and evolution of the Badain Jaran Desert, North China, as revealed by a drill core from the desert centre and by geological survey. Palaeogeogr Palaeoclimatol Palaeoecol 426: 139–158. doi:10.1016/j.palaeo.2015.03.011

Wilson JS, Pitts JP (2012) Identifying Pleistocene refugia in north American cold deserts using phylogeographic analyses and ecological niche modelling. Divers Distrib 18(11): 1139–1152. doi:10.1111/j.1472-4642.2012.00902.x

Yang Z, Rannala B (2014) Unguided species delimitation using DNA sequence data from multiple loci. Mol Biol Evol 31(12): 3125–3135. doi:10.1093/molbev/msu279

Zhang RZ (2011) Zoogeography of China. Science Press, Beijing

Zhang LS (2012) Palaeogeography of China-the formation of Chinese natural Evironment. Science Press, Beijing

Zhang ZS, Guo ZT (2005) Spatial character reconstruction of different periods in Oligocene and Miocene. Quaternary Sciences 25(04): 523–530

Zhang YJ, Stock M, Zhang P, Wang XL, Zhou H, Qu LH (2008) Phylogeography of a widespread terrestrial vertebrate in a barely-studied Palearctic region: Green toads (*Bufo viridis* subgroup) indicate glacial refugia in eastern Central Asia. Genetica 134(3): 353–365. doi:10.1007/s10709-008-9243-0

Zhang Q, Xia L, Kimura Y, Shenbrot G, Zhang ZQ, Ge DY, Yang QS (2013) Tracing the origin and diversification of Dipodoidea (order: Rodentia): Evidence from fossil record and molecular phylogeny. Evol Biol 40(1): 32–44. doi:10.1007/s11692-012-9167-6

Zhou WW, Zhang BL, Chen HM, Jin JQ, Yang JX, Wang YY, Jiang K, Murphy RW, Zhang YP, Che J (2014) DNA barcodes and species distribution models evaluate threats of global climate changes to genetic diversity: A case study from *Nanorana parkeri* (Anura: Dicroglossidae). PLoS One 9(8): e103899. doi:10.1371/journal.pone.0103899