Phylogeography of the desert scorpion *Mesobuthus mongolicus* illuminates a route out of Central Asia and Junggar Basin as a corridor for faunal exchange between Central and East Asia

Cheng-Min Shi^{a,*}, Xue-Shu Zhang^b, Lin Liu^b, Ya-Jie Ji^b, and De-Xing Zhang^{b,c*}

^aState Key Laboratory of North China Crop Improvement and Regulation, College of Plant Protection, Hebei Agricultural University, Baoding 071000, China.

^bState Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China.

^cUniversity of Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100049, China.

*Address correspondence to Cheng-Min Shi: shichengmin@hebau.edu.cn or De-Xing Zhang: dxzhang@ioz.ac.cn

Running head: Phylogeography of *Mesobuthus mongolicus*

Abstract

A comprehensive understanding of the ecological, evolutionary and genetic consequences of climate changes requires integration of information from different geographic regions and ecosystems. However, a clear knowledge gap exists in arid biota of Asian drylands, where climate change and biological evolution demonstrate high regional specificity. We report here our phylogeographic study on the desert scorpion *Mesobuthus mongolicus*, attempting to narrow the gap. Range wide samples for one mitochondrial gene and three nuclear loci were subjected to phylogenetic inferences, molecular dating, Bayesian phylogeographic and demographic reconstructions, and population genetic analyses. Ecological niche modeling was employed to generate models of potential distribution range at present and during the last glacial maximum (LGM). Our results revealed that *M. mongolicus* represents a single coherent lineage that diverged with its most closely related lineage from Central Asia about 1.36 Ma and underwent radiation ever since. Bayesian phylogeographic reconstruction suggested the ancestral population of *M. mongolicus* dispersed gradually eastward through the Junggar Basin to the Gobi region during the late Pleistocene, strengthening the view that the Junggar Basin has constituted an important corridor for faunal exchange between Central Asian and East Asia. We found that the desert scorpion underwent demographic expansion while experiencing range contraction during the LGM. Development of sandy deserts in northwest China incurred by Eurasian aridification might have opened up substantial empty niches that sustained population expansion during the LGM. Our results suggested that organisms' ecological adaptations have acted as an important determinant for their phylogeographic and demographic responses to past climate changes.

Keywords: climate oscillations, demographic expansion, drylands, *Mesobuthus*, phylogeography

Geographic distribution of a species is an outcome of interplay between environmental challenges and its biological adaptation. Given a certain environmental change, it can either pose a threat or provide an opportunity depending on species' demographic, physiological and behavior adaptational profiles (Garcia et al. 2014). Thus, species might respond differently to climatic changes in an individualistic way (Ackerly et al. 2010; Anadón et al. 2015; Hewitt 2000; Willis et al. 2004). Besides, historical climate changes were expressed variably across geographic regions (Hewitt, 2000), and their impacts were unfolded in multiple dimensions across geographic locations (Garcia et al. 2014). These features make comprehension of the biological and ecological impacts of climate changes a very complex issue that requires to integrate information from a diverse of organisms, ecosystems, geographic locations and climate regimes.

However, studies on such topics are focused predominantly on Europe and North America and essentially on the impact of glaciations (Keppel et al. 2012). In those regions, it is generally well acknowledged that the glacial cycles have played a major role in shaping geographic distributions and reshuffling genetic variations of plants and animals (Hewitt 2000). On the contrary, the arid regions and the role of aridification, a more widespread coeval of the Quaternary glaciation, remain to be fully addressed (Byrne et al. 2008, 2018; Shi et al. 2013).

Drylands, including deserts and semi-deserts, cover about 40% of the Earth's terrestrial surface, stretching particularly widely in mid latitude Eurasia (Middleton and Sternberg 2013). Those regions encompass a wide variety of physical terrains. However, what all drylands have in common is aridity, expressed as characteristically scarce precipitation and extremely high evaporation. Such climatic conditions have shaped some of most sensitive and fragile desert ecosystems on the Earth (Austin 2011). Since 22 million years ago, aridification has been a surrogate of climatic changes and profoundly shaped the evolution of biota in the Asian interior (Bougeois et al. 2018; Guo et al. 2002). It has been suggested that the phylogeographical effects of aridification in the subtropic hot deserts mirror or contrast, depending on the ecological adaptation of studied species, the glaciation in the Northern Hemisphere (Fujita et al. 2010). However, this conjecture remains to be tested in the cold deserts in the temperate regions, particularly the drylands in Asian interior.

An expansion-contraction (EC) model has been proposed to describe the phylogeographic dynamics of biota in response to the Pleistocene glacial oscillations (Provan and Bennet 2008). EC model has stood extensively empirical corroboration with temperate biota in Europe and North America. For temperate species, the advance and retreat of ice sheets imposed a consequent parallel contraction-expansion of species' range and population size (Hewitt 2004; Provan and Bennett 2008; Salvi et al. 2014; Schmitt 2007). On the contrary, for cold-adapted taxa, an inverse model has been proposed (Galbreath et al. 2009; Willis et al. 2004). Evidence is accumulating that EC model is also applicable to subtropical arid ecosystems although the underlying driving climatic factor might differ. In those hot deserts, precipitation rather than temperature, is the limiting resource (Noy-Meir 1973), and has exerted more impact on phylogeographic history (Cosacov et al. 2013). For example, Australia's arid-adapted Grey

Butcherbird (*Cracticus torquatus*) experienced population expansions during Pleistocene glacial maxima whereas population of its mesic-adapted relatives remained stable (Kearns et al. 2014).

Here we test the EC model in the Asian cold deserts including the Gobi where both temperature and aridity pose great challenges to biota, using the desert scorpion *Mesobuthus mongolicus* (Birula, 1911) as the studying organism. This scorpion was formerly known as *M. eupeus mongolicus* and has recently been upgraded to an independent species (Kovářik 2019). It represents the easternmost member of the *M. eupeus* complex (Fet et al. 2018; Fet et al. 2000; Shi et al. 2015). Previously, we have shown that the ancestral distribution area of the *M. eupeus* complex is most likely in Anatolia (Shi et al. 2013). Gobi is the largest desert of Asia, covering huge parts of northern China and of southern Mongolia (ca. 1.3 million km²). It is located far north than other temperate deserts on a plateau more than 900 m above sea level. As a consequence, the temperature in Gobi can be extremely hot (>50 °C) in summer and extremely cold (-49 °C) in winter (Pfeiffer et al. 2003). Locked by rain shadows created by the lofty Qinghai-Tibet Plateau and the Tianshan and Altai mountains, Gobi receives very scarce rainfall but endures extremely heavy evaporation. On such an ‘edge of life’, however, *M. mongolicus* (Mmon hereafter) thrives (Shi et al. 2007, 2015), providing an ideal model to bridge the knowledge gap in dryland phylogeography. The Gobi populations of Mmon has been proposed as a product of out-of-Central Asia move of its ancestors during the late Pleistocene (Shi et al. 2013). However, the exact route through which Mmon migrated to the Gobi remains to be defined. In the present study, we have tested the EC model by combining genetic analysis and ecological niche modelling of Mmon, with the aim to explore (1) the possible route through which Mmon dispersed out of Central Asia, (2) its demographic pattern casted by past climate changes, and (3) the climatic context that has shaped its phylogeography.

Materials and Methods

Sample collection and genetic data

The natural distributional range of Mmon has been determined through systematic field survey with the aid of species distribution modeling (Shi et al. 2015). We sampled 338 individuals of Mmon from 51 sites, including 41 sites in China (C₁-C₄₁, *n* = 274), 8 sites in Mongolia (M₁-M₈, *n* = 62), one site each in Tajikistan (TJ1, *n* = 1) and Kazakhstan (KZ1, *n* = 1) (Figure 1). The sample sizes ranged from 1 to 16 individuals. Of those, 44 individuals were from Shi *et al.* (2013), the remaining 294 individuals were sampled in this study. We also included 12 individuals *M. przewalskii* (Zhang et al. 2020), and 10 individuals of *M. martensii* (Shi et al. 2013) in the phylogenetic analysis. These two species occur parapatrically with Mmon.

Genomic DNA was extracted from preserved tissues using a modified phenol-chloroform extraction procedure (Zhang and Hewitt 1998). All individuals were sequenced for the mitochondrial cytochrome C oxidase subunit I (mtCOI) gene and a subsample were sequenced for three nuclear loci, viz: nuclear protein kinase (PK), Serin-type endopeptidase (STE) and Serin proteinase inhibitor (Spn2) genes. The primers, PCR profiles and sequencing protocols followed Shi *et al.* (Shi et al. 2013). All specimens and

DNA extracts were deposited at the Laboratory of Molecular Ecology and Evolution, Institute of Zoology (MEE-IOZ), Chinese Academy of Sciences, Beijing. The unique haplotypes have been deposited in GenBank under accession numbers xxxxxxxx-xxxxxxx.

Phylogenetic analyses and divergent time estimate

Sequences were aligned using MUSCLE (Edgar 2004). Phylogenetic analyses were performed using both maximum likelihood (ML) and Bayesian methods. The best-fit model of DNA evolution was selected using Modeltest 3.7 program (Posada and Crandall 1998) according to the Akaike information criterion (AIC). ML analysis was carried out using PHYML 3.0 (Guindon et al. 2010) with SPR tree search operation. Bootstrap with 1000 replicates was performed to assess topological robustness. Bayesian analysis was carried out with MrBayes 3.2 (Ronquist et al. 2012). Analyses were initiated with random starting trees and run for 10^6 generations with four Markov chains employed. Trees were sampled every 100 generations and the ‘temperature’ parameter was set to 0.2. The first 30% trees were discarded as burn-in after checking for stationary and the convergence. Median-joining network (Bandelt et al. 1999) was constructed using the Network 4.6 .1.2 (<http://www.fluxus-engineering.com/>) to visualize intraspecific relationships among distinct haplotypes.

Divergence time was estimated with a relaxed molecular clock approach implemented in BEAST 1.10.4 (Suchard et al. 2018). We combined the most diverged mtCOI haplotypes revealed here with our previous alignment into a new data set with 53 sequences and estimated divergent times with the same calibrating method (Shi et al. 2013). The rate change was explicitly modeled using uncorrelated lognormal distribution across tree. Using a GTR + Γ model and a Yule process for speciation prior, we performed 100 million Markov chain Monte Carlo (MCMC) searches, sampling once every 1000 generations. Stationary and the convergence of the MCMC chains were checked with Tracer 1.7 (Rambaut et al. 2018). Maximum clade credibility (MCC) tree, posteriors, means and 95% highest posterior densities (HPDs) of ages of nodes were identified and annotated using TreeAnnotator 1.10 (Suchard et al. 2018).

Phylogeographic diffusion analysis

Current and predicted suitable distribution area of Mmon stretches continuously across the entire studied region (Shi et al. 2015). There are no obvious potential barriers that can be enrolled to define discrete areas, and thus it is arbitrary to divide species range into discrete operational areas *a priori*. Therefore, we inferred spatial diffusion of genealogies using a continuous Bayesian phylogeographic method which models geographic spread of genealogies across spaces with continuous diffusion processes conditioned on sampling location defined as bivariate trait representing latitude and longitude for each accession (Lemey et al. 2010). This analysis was only based on the mtCOI data due to limited spatial sampling and low genetic variation of the nuclear genes. We first selected the best fit diffusion model by performing marginal likelihood estimation using generalized stepping-stone sampling on three available relaxed

random walk models, as well as the time-homogeneous Brownian motion process. The Bayesian phylogeographic diffusion analysis was performed using BEAST v. 1.10.4 (Suchard et al. 2018). MCMC were run for 50 million generations with trees and parameters sampled every 5,000 generations. MCMC samples were summarized using TreeAnnotator. The results of phylogeographic reconstruction were visualized in Spred3 (Bielejec et al. 2016).

Demographic inferences

Historical changes in the effective population size were assessed using two approaches. First, we used summary statistic tests, D (Tajima 1989), F_s (Fu 1997) and R_2 (Ramos-Onsins and Rozas 2002). Tajima's D test is based on the number of segregating sites, and Fu's F_s test uses information from the haplotype distribution, while Ramos-Onsins and Rozas's R_2 test uses information of the mutation frequency. These statistics have high power to detect population expansion for non-recombining regions of the genome under a variety of different circumstances, either when sample size is large (F_s) or when sample sizes is small (D), and the number of segregating sites is low (R_2). The significance of each test was assessed by generating null distribution from 10,000 coalescent simulations in DnaSP version 5.10.01 (Librado and Rozas 2009). Significantly large negative D and F_s values and significantly positive R_2 value were taken as evidence of population expansion. Second, demographic changes over time were explored by coalescent-based Bayesian skyline plot (BSP) using BEAST 1.10.4 (Suchard et al. 2018). Unlike above tests, which are based on summary statistics, BSP makes use of all the historical information contained within a sample of DNA sequences from a given population. The time estimated from molecular dating was used to calibrate BSP analyses. We ran the analysis for 5×10^8 iterations, treating the first 10% iterations as burn-in, and then sampled genealogies and model parameters every 5, 000 generations for subsequent analysis.

Ecological Niche Modeling

A total of 89 occurrence records for Mmon were used for ecological niche modeling (ENM) with MaxEnt version 3.3.3k (Phillips et al. 2006). We removed localities that were within distances of 30 km to each another to reduce the effect of sampling bias. We used the bioclimatic variables for both the present and past (21 kyr B.P., based on MIRCO models) at 2.5 arc-minute resolution (c. 5×5 km) in the WorldClim data set (Hijmans et al. 2005). The present bioclimatic layers were masked to region spanning from 35.8 to 55.4 °N and from 76.4 to 122.5 °E, to avoid sampling unrealistic background data and thus inflating the strength of predictions. The pairwise Pearson's correlations between the climatic variables were used to restrict climatic variables. We removed highly correlated ($|r| \geq 0.80$) climatic variables before ENM.

MaxEnt was run with a convergence threshold of 10^{-5} and maximum number of iterations of 10,000 with 25% of localities for model testing for 10 replicates with subsample. The model was projected to past conditions with clamping. Model performance was assessed via the area under the ROC (receiver operating characteristic) curve (AUC) statistic and the importance of variables was assessed by jackknife

tests. AUCs > 0.75 are considered adequate for species distribution modeling applications (Fielding & Bell, 1997; Pearce & Ferrier, 2000). We employed two threshold rules, the maximum training sensitivity plus specificity threshold (MTSS), and the equate entropy of thresholded and original distributions logistic threshold (ENT), for converting continuous models to binary predictions. The MTSS is considered to most accurately predict presence/absence (Jiménez-Valverde and Lobo 2007). The ENT is known significantly over-predicts species ranges (Worth et al. 2014).

Results

Phylogenetic relationships and divergence times

The total 338 mtCOI sequences collapsed into 62 unique haplotypes, of which one haplotype was from Tajikistan (TJ01), one haplotype was from Kazakhstan (KZ01), and all other 60 haplotypes were from northwest China and south Mongolia. Combining with 12 samples of *M. przewalskii*, 10 samples of *M. martensii*, 6 representatives for another three species of *M. eupeus* complex and two outgroups, the final dataset includes 92 mtCOI haplotype sequences. The results of ML analysis and Bayesian inference were highly congruent. The ML phylogeny inferred from mtCOI dataset is shown in Figure 2A. All samples of Mmon from China and Mongolia clustered within an unresolved star clade which coalesced with TJ01 from Tajikistan with strong supports (95/1.00 for ML/Bayesian, respectively). We recognize this clade as Mmon. Within the Mmon clade, all haplotypes except one from the Junggar Basin (Xinjiang) formed a sub-lineage but was not strongly supported. This result was mirrored by the haplotype network analysis. Haplotypes from the Junggar Basin were closely related to those from the Gobi region but no haplotype was shared between two regions (Figure 2B). The Mmon clade coalesced with KZ01 and then *M. thersites* and *M. phillipsi*, forming a monophyletic branch (71/0.99) including exclusively members of the *M. eupeus* species complex. Monophyly for both *M. przewalskii* and *M. martensii* was supported. These two species were more closely related each other than to any member of the *M. eupeus* complex.

The concatenated matrix for three nuclear loci (PK, STE and Spn2) included 33 samples. The results of phylogenetic analyses were congruent with the mitochondrial phylogeny as far as the interspecies relationship was concerned (Figure 2C). However, the nuclear phylogeny provided more resolved relationships within Mmon. Two moderately supported subclades were observed. One subclade includes three samples from the Junggar Basin and TJ01 from Tajikistan. The other subclade includes all samples from the Gobi region and three clustered samples from the Junggar Basin.

The results of divergence time estimation were summarized in Table 1. The time of the most recent common ancestor (T_{MRC}) of Mmon was estimated to be 1.36 Ma (95% HPD, 0.78–1.97 Ma) and the Junggar-Gobi lineage coalesced at 0.54 Ma (0.30–0.80 Ma). The divergence time between Mmon and KZ01 is 3.37 Ma (2.12–4.54 Ma), while that between *M. przewalskii* and *M. martensii* is 6.27 Ma (4.35–8.37 Ma). A further split within *M. przewalskii* occurred at 1.00 Ma (0.56–1.51 Ma). Divergence between the *M. eupeus* complex and the *M. przewalskii*-*M. martensii* clade occurred around 10.56 Ma (8.33–12.70 Ma).

The spatiotemporal diffusion of *Mmon*

A Bayesian phylogeographic reconstruction of ancestral distribution and spatial diffusion of genealogies through time supports that *Mmon* was originated in Central Asia and dispersed eastward step by step, and finally to the Gobi regions. Three snapshots illustrating the major phylogeographic dispersal events are shown in Figure 3. The most likely ancestral distribution of the last most recent common ancestors of *Mmon* located in an area to the northwest of the Tianshan, somewhere near the boundary between Tajikistan and Kazakhstan. The ancestral population moved eastward and reached the Junggar Basin around 0.8 Ma, and then underwent regional dispersion within the Junggar Basin. Around 0.4 Ma, *Mmon* dispersed further east to the southern part of the Gobi, followed by another regional expansion in areas near the Tengger Desert and the Badan Jaran Desert, striding Gansu, Inner Mongolia and Ningxia. It was a very recent event (<0.2 Ma) that *Mmon* population expanded northward to the Gobi region of southern Mongolia. During this time, a northward expansion event was also occurred in the Junggar basin which gave rise to the Altay population. It appeared that regional population expansions have persisted until the present day, which gives rise to current trans-Gobi distribution for this species.

Demographic history

All the values for Fu's F_s were negative and significant ($P < 0.01$), and all the values of Ramos-Onsins and Rozas' R_2 were small (< 0.12), positive and significant ($P < 0.05$) for populations of *Mmon* (Table 2). The values of Tajima's D for all except the Junggar population were negative and significant ($P < 0.01$). All these tests for *M. przewalskii* were not significant ($P > 0.3$, data not shown). The variations of effective population size through time are shown by Bayesian skyline plots in Figure 4. For *Mmon*, the curve of BSP can be visually divided in to three phases. Before 35 ka, *Mmon* population were very stable. From 34 to 14 ka, *Mmon* population size increased sharply, even during the LGM around 21 ka. Since 14 ka until to present, the BSP curve of *Mmon* gradually plateaued off (Figure 4A). The population size of *M. przewalskii* increased slightly since 400 kya and leveled off approaching the present day (Figure 4B).

Temporal changes of suitable distributional areas

After removing highly correlated ($|r| \geq 0.80$) climatic variables, six climatic variables (BIO1, annual mean temperature; BIO7, temperature annual range; BIO9, mean temperature of driest quarter; BIO12, annual precipitation; BIO15, precipitation seasonality; BIO19, precipitation of coldest quarter) were used for final ecological modeling. These variables seem to be biologically meaningful for *Mmon* considering its ecology and the climatic regime of the studied regions (Shi et al. 2007). The distributional models generated with these six climatic variables maintained high AUC statistics (average 0.928 ± 0.023). The potential distribution of *Mmon* averaged over 10 models for the present climate are shown on Figure 5A. It encompassed the entire Junggar Basin and the Gobi region, with a narrow belt along the Hexi Corridor connecting these major areas. The northern rim of the Tarim Basin was also suitable for distribution of

Mmon. However, both the Tianshan and Altai mountains were obviously unsuitable. Figure 5B shows the Palaeo-projection of ENMs at the LGM. It indicated sharp contraction of suitable range at this time, resulting in two major disjointed areas in the Junggar Basin and to the south of Gobi, respectively. In those regions, extensive southward retreat of potential niche occurred in the Gobi region, with the northern Mongolian Gobi becoming completely unsuitable and the Badian Jaran-Tengger sand deserts being the center of distribution. In the western part, the suitable area shrank towards the heart of the Junggar Basin. Of the 6 climate variables used, annual precipitation (BIO12) contributed to the MaxEnt models most (with a mean relative contribution of 47.2 %), followed by annual mean temperature (BIO1, 31.1%), and mean temperature of driest quarter (BIO9, 16.4%). The three remaining variables collectively contributed 5.4% to the model.

Discussion

Past climatic oscillations have triggered recurrent reorganization of biota at all levels and scales, from genes to ecosystems (Davis and Shaw 2001; Hewitt 2000). In response to the climatic changes, populations must migrate to remain within their ecological niche or adapt to tolerate the challenges in order to survive and thrive (Davis and Shaw 2001; de Lafontaine et al. 2018; Fernandez et al. 2021). Otherwise, they will risk extinction. Scorpions of *Mesobuthus* (*sensu* Gantenbein et al. 2003) faithfully followed this general rule, thus emerging as an excellent model for studying biogeographical evolution in drylands of Asian interior (Graham et al. 2019; Shi et al. 2013). Since the late Miocene, this group has undergone extensive diversification and biogeographic evolution driven by regional tectonics (Mirshamsi et al. 2010; Mirshamsi et al. 2011) and climate changes (Shi et al. 2013), which gave rise to more than two dozen species widely spread in mid latitude Eurasia with Central Asia being center of species diversity (Fet et al. 2021; Fet et al. 2018; Fet et al. 2000; Graham et al. 2019). Currently, at least five species of *Mesobuthus* occur in China which was hypothesized as an outcome of multiple ‘Out of Central Asia’ evolution of the ancestral species (Shi et al. 2013). The Gobi population of Mmon was suggested to have undergone very recent out-of-Central Asia move in late Pleistocene (Shi et al. 2013). In the present study, we first define the route via which ancestral population of Mmon dispersed out of Central Asia and reached the Gobi, then tested the expansion-contraction model that might have incurred by climate-driven dispersal events, and finally explored the environmental context that has sustained such phylogeographic history.

The Junggar Basin as the route out of Central Asia for the desert scorpion

Our study provides several lines of evidence that Mmon dispersed out of Central Asia and colonized the Gobi via the Junggar Basin along the northern side of the Tianshan mountains. Previously, we have shown that the ancestral distribution area of the *M. eupeus* complex was most likely in Anatolia and the ancestral population of Mmon arose somewhere in Central Asia (Shi et al. 2013). In the context of regional tectonic history and current topographic configuration such as the high mountain ranges (*i.e.*

Tianshan and Pamir), it was possible for ancestral Mmon to disperse to the Gobi either through the Junggar Basin to the north of Tianshan (Junggar route) or through the Tarim Basin to the south of Tianshan (Tarim route), or by taking both routes simultaneously. Our results suggest that the Junggar route was the most likely one.

With extensive samples from China and Mongolia, we identified a single geographically coherent Mmon lineage which coalesced around 0.54 Ma and diverged from its most closely related lineage in Central Asia (TJ01) around 1.36 Ma (Table 1). Given ancient uplift of the Tianshan mountains forming a formidable geographical barrier (Figure 5), we would expect a much deeper genetic divergence than we observed here between the Central Asian population (*e.g.* TJ01) and the Gobi populations, if the desert scorpion dispersed out of the Central Asia along the Tarim route. During the late Miocene, most part of the Tianshan mountains have reached a critical height to be an effective geographical barrier (Yin 2010). The single lowland gateway connecting Central Asia and the Tarim Basin closed at 5.3 Ma due to the collision of the Tianshan orogen and the Pamir salient (Sun et al. 2015). Thus, if Mmon had dispersed via the Tarim route, it would imply a divergence date of around 5.3 Ma between the Central Asia population and the Gobi population. This date is more than three times earlier than the date we have estimated here (1.36 Ma, Table 1). Thus, the Tarim route for Mmon dispersal out of Central Asia can be largely excluded.

Indeed, Bayesian phylogeographic reconstruction clearly supported the Junggar route out of Central Asia for Mmon (Figure 3). During the past 1.36 million years, Mmon has been on an active eastward move. The ancestral population entered the Junggar Basin around 0.8 Ma and reached the Gobi region round 0.4 Ma. We noticed that such a colonization route from Central Asia was not limited to the scorpion species we have studied. Similar phylogeographic patterns have been reported in the agamid lizards (*Phrynocephalus* sp.) (Melville et al. 2009), and the rapid racerunner (*Eremias velox*) (Liu et al. 2019), although the times of colonization might differ. These evidences suggest that the Junggar Basin have constituted an important corridor for fauna exchange between Central Asia and the Gobi regions.

Demographic expansion of the desert scorpion decoupled from range expansion during the LGM

A typical response of temperate species to past climate changes is demographic fluctuations. The star-shaped genealogy (Figure 2AC), the excess of rare mutations and deficit of intermediate frequency mutation (Fig 2B), and the significant negative Tajima's *D* and Fu's *F_s* as well as significant small *R₂* (Table 3), all supported population expansion in Mmon. The curve of BSP indicated that Mmon population underwent exponential growth since about 35 ka and throughout the LGM (Figure 3A). Such a glacial maximum expansion phenomenon is in sharp contrast to the paradigmatic interglacial-expansion glacial-contraction model corroborated for temperate flora and fauna of North America and Europe (Hewitt 2004; Schmitt 2007). It is also inconsistent with the scenario for the mesic biota in arid regions of the other parts of the world, in which population contraction occurred during the LGM with the effect of aridification mirrored that of glaciation (Fujita et al. 2010). However, such a demographic pattern

appears to be characteristic of arid-adapted species in northwest China. For example, population expansions during the LGM have been detected in the desert-dwelling gerbil *Meriones meridianus* (Wang et al. 2013), the desert plant species of *Lagochilus* (Meng and Zhang 2013) and *Zygophyllum xanthoxylon* (Shi and Zhang 2015). Glacial demographic expansions also occurred in arid-adapted species from other regions of the world, for instance, the grey butcherbird *Cracticus torquatus* (Kearns et al. 2014) from arid woodland of southern Australia, the lizard *Liolaemus darwini* in the Monte Desert of South America (Camargo et al. 2013), and three sympatric beetle species (*Iliotona beyeri*, *Carcinops gilensis* and *Belonuchus* sp.) from the Sonoran Desert of North America (Pfeiler et al. 2013). Our results, along with these cases, substantiate that organisms' ecological adaptations were an important determinant to their demographic response to past climate changes.

It has been repeatedly demonstrated in a plethora of animals and plants that demographic expansion was often accompanied by range expansion. However, Mmon appeared to be an exception. We observed that demographic expansion of Mmon was accompanied by range contraction. While Mmon experienced population expansion, ENM indicated that its geographical distribution was drastically reduced during the LGM. That is, the suitable climate niche shrank toward the southern Junggar Basin and the southern Gobi (Figure 5). Nevertheless, such a scenario was reasonable when considering collectively the regional climate context and the scorpion ecology. Firstly, climate change in these regions seemed moderate at the LGM. Only limited glacial expansion occurred in Tianshan and other nearby high mountains during the LGM (Lifton et al. 2014). Secondly, Mmon is an arid-adapted species, and thus aridification associated with glaciation would exert limited stress to its persistence. On the contrary, the consequential expansion of desert environments might have opened up an array of new habitats that became accessible to Mmon. Besides, the extent of sand dune areas was expanded during glacial periods (Dong et al. 2013; Lu et al. 2013), which would provide Mmon suitable matrix for burrowing, by which scorpions could effectively buffer the climate extremes. As a consequence, although the overall size of their niche area was reduced during the LGM, the localized effective niche spaces might have expanded with the development of the desert ecosystems. Third, given that Mmon is the only scorpion species that has colonized the Gobi region rather recently (Shi et al. 2013) and that scorpions can maintain extremely high biomass in arid communities (Polis and Yamashita 1991), it could be expected that the population density would be far below the carrying capacity of the habitats even at the LGM, allowing proliferation of Mmon populations. In summary, the unique ecological adaptation of the desert scorpion and the distinct regional climate regime should have collectively molded the observed demographic pattern of Mmon in Asian drylands.

The environmental context for phylogeographic evolution in Asian drylands

The evolutionary history of the desert scorpion Mmon reconciled well with the history of environmental changes in Asia drylands. The single geographically coherent lineage from China and Mongolia diverged from its most closely related lineage in Tajikistan around 1.36 Ma (0.78-1.97 Ma). Around this time,

Central Asia become increasingly arid (Ding et al. 2002; Yang and Ding 2006) and northern China experienced an abrupt transition to sharply intensified winter monsoon (Ding et al. 1994; Liu et al. 1999), accompanied by expansion of desert environment (Ding et al. 2005). Study on the loess in Tajikistan also showed an increased aridity in Central Asia since 0.85 Ma (Ding et al. 2002; Yang and Ding 2006). Eolian loess deposited on the northern slope of the Tianshan indicated that the extremely dry climatic conditions was present in the Junggar Basin at least by 0.8 Ma and that the desert there experienced two dramatic periods of expansions at ~0.65 Ma and 0.5 Ma, respectively (Fang et al. 2002). These environmental changes might have provided the desert scorpion *Mmon* an opportunity to expand further eastward out of the Junggar Basin. To more eastern region, Badain Jaran and Tengger sand deserts were in place around 0.85 Ma (Guan et al. 2011; Li et al. 2014). The arid desert landscape occurred on the alluvial plain of the Gobi-Altai mountains no later than 0.42 Ma (Lü et al. 2010). The sudden radiation of *Mmon* and its out of Central Asia dispersal (around 0.54 Ma, 0.3-0.8 Ma) was in the frame of these aridification events. These stepwise expansion of desert environment (Ding et al. 2005) should have opened up suitable habitats for *Mmon* to disperse eastward and finally reached the Gobi region. As revealed by hindcast ENM, large suitable areas existed both in the Junggar Basin and in the Gobi region although become disjunct at LGM (Figure 5B). Such large refugial regions would have sustained demographic expansion even during the LGM (Fig 4).

In summary, by revealing a specific phylogeographic history of the desert scorpion *M. mongolicus*, which was relevant to its biology and reconciled to regional climate change regime, we have demonstrated that (1) the eastward dispersal of the desert scorpion out of Central Asia to the Gobi was driven by intensified aridification events, and most likely through the Junggar Basin which have constituted an important corridor for fauna exchange between Central Asia and the Gobi regions, (2) the demographic expansion of the desert scorpion was accompanied by range contraction during the LGM, and thus (3) an organism's ecological adaptation were an important determinant to their phylogeographic and demographic response to past climate changes. Future work directing to understanding the crucial limiting factors of species range from different geography will provide major scaffolds for integrating phylogeographic knowledge across spaces and times.

Acknowledgements

We are grateful to the Plant Protection Research Institute, Mongolia, for logistic supports on field surveys. We thank Peng He, Hong-Bin Liang, Ke-Qing Song, Zhiyuan Su, Duo-Hong Wang, Zhi-Liang Wang for collecting scorpions.

Fundings

This study was supported by the National Natural Science Foundation of China (grant Nos. 31772435, 31370041). C-M Shi is also supported by a starting fund from Hebei Agricultural University and the State Key Laboratory of North China Crop Improvement and Regulation (YJ2020028).

Author contributions

D-X.Z. and C-M.S. conceived and designed the project; C-M.S., X-S.Z., L.L and Y-J.J. collected the molecular data; C-M.S. performed the field work and analyzed data; C-M.S. and D-X.Z. wrote the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- Ackerly DD, Loarie SR, Cornwell WK, Weiss SB, Hamilton H et al., 2010. The geography of climate change: Implications for conservation biogeography. *Diver Distrib* **16**:476-487.
- Anadón JD, Graciá E, Botella F, Giménez A, Fahd S et al., 2015. Individualistic response to past climate changes: Niche differentiation promotes diverging Quaternary range dynamics in the subspecies of *Testudo graeca*. *Ecography* **38**:956-966.
- Austin AT, 2011. Has water limited our imagination for aridland biogeochemistry? *Trends Ecol Evol* **26**:229-235.
- Bandelt HJ, Forster P, Röhl A, 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**:37-48.
- Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A et al., 2016. Spread3: Interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol Biol Evol* **33**:2167-2169.
- Bougeois L, Dupont-Nivet G, de Rafélis M, Tindall JC, Proust J-N et al., 2018. Asian monsoons and aridification response to Paleogene sea retreat and Neogene westerly shielding indicated by seasonality in Paratethys oysters. *Earth Planet Sci Lett* **485**:99-110.
- Byrne M, Yeates D, Joseph L, Kearney M, Bowler M et al., 2008. Birth of a biome: Insights into the assembly and maintenance of the Australian arid zone biota. *Mol Ecol* **17**:4398-4417.
- Byrne M, Joseph L, Yeates DK, Roberts JD, Edwards D, 2018. Evolutionary history. In H. Lambers (Ed.), *On the ecology of Australia's arid zone* (pp. 45–75). Springer International Publishing.
- Camargo A, Werneck FP, Morando M, Sites Jr JW, Avila LJ, 2013. Quaternary range and demographic expansion of *Liolaemus darwini* (Squamata: Liolaemidae) in the monte desert of Central Argentina using Bayesian phylogeography and ecological niche modelling. *Mol Ecol* **22**:4038-4054.
- Cosacov A, Johnson LA, Paiaro V, Cocucci AA, Córdoba FE, 2013. Precipitation rather than temperature influenced the phylogeography of the endemic shrub *Anarthrophyllum desideratum* in the patagonian steppe. *J Biogeogr* **40**:168-182.
- Davis M, Shaw R. 2001 Range shifts and adaptive responses to quaternary climate change. *Science* **292**:673-679.
- de Lafontaine G, Napier JD, Petit RJ, Hu FS, 2018. Invoking adaptation to decipher the genetic legacy of past climate change. *Ecology* **99**:1530-1546.
- Ding Z, Yu Z, Rutter NW, Liu T, 1994. Towards an orbital time scale for Chinese loess deposits. *Quat Sci Rev* **13**:39-70.
- Ding ZL, Derbyshire E, Yang SL, Sun JM, Liu TS, 2005. Stepwise expansion of desert environment across northern China in the past 3.5 Ma and implications for monsoon evolution. *Earth Planet Sci Lett* **237**:45-55.
- Ding ZL, Ranov V, Yang SL, Finaev A, Han JM et al., 2002. The loess record in southern Tajikistan and correlation with Chinese loess. *Earth Planet Sci Lett* **200**:387-400.
- Dong Z, Qian G, Lv P, Hu G, 2013. Investigation of the sand sea with the tallest dunes on Earth: China's Badain Jaran sand sea. *Earth Sci Rev* **120**:20-39.
- Edgar RC, 2004. Muscle: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792-1797.

- Fang X, Shi Z, Yang S, Yan M, Li J et al., 2002. Loess in the Tian Shan and its implications for the development of the Gurbantunggut Desert and drying of northern Xinjiang. *Chin Sci Bull* **47**:1381-1387.
- Fernandez MC, Hu FS, Gavin DG, de Lafontaine G, Heath KD, 2021. A tale of two conifers: Migration across a dispersal barrier outpaced regional expansion from refugia. *J Biogeogr* **48**:2133-2143.
- Fet V, Kovařík F, Gantenbein B, Graham MR, 2021. Three new species of *Olivierus* (Scorpiones: Buthidae) from Kazakhstan and Uzbekistan. *Zootaxa* **5006**:54-72.
- Fet V, Kovařík F, Gantenbein B, Kaiser R, Stewart A et al., 2018. Revision of the *Mesobuthus caucasicus* complex from Central Asia, with descriptions of six new species (Scorpiones: Buthidae). *Euscorpius* **255**:1-77.
- Fet V, Sissom W, Lowe G, Braunwalder M, 2000. Catalog of the Scorpions of the World (1758–1998). New York: The New York Entomological Society.
- Fu YX, 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**:915-925.
- Fujita MK, McGuire JA, Donnellan SC, Moritz C, 2010. Diversification and persistence at the arid–monsoonal interface: Australia-wide biogeography of the Bynoe’s gecko (*Heteronotia binoei*; Gekkonidae). *Evolution* **64**:2293-2314.
- Galbreath KE, Hafner DJ, Zamudio KR, 2009. When cold is better: Climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, *Ochotona princeps*). *Evolution* **63**:2848-2863.
- Gantenbein B, Fet V, Gromov AV, 2003. The first DNA phylogeny of four species of *Mesobuthus* (Scorpiones, buthidae) from Eurasia. *J Arachnol* **31**:412-420, 419.
- Garcia RA, Cabeza M, Rahbek C, Araújo MB, 2014. Multiple dimensions of climate change and their implications for biodiversity. *Science* **344**:1247579.
- Graham MR, Myers EA, Kaiser RC, Fet V, 2019. Cryptic species and co-diversification in sand scorpions from the Karakum and Kyzylkum deserts of Central Asia. *Zool Scr* **48**:801-812.
- Guan Q, Pan B, Li N, Zhang J, Xue L, 2011. Timing and significance of the initiation of present day deserts in the northeastern Hexi Corridor, China. *Palaeogeogr Palaeoclimatol Palaeoecol* **306**:70-74.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W et al., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of phyml 3.0. *Syst Biol* **59**:307-321.
- Guo ZT, Ruddiman WF, Hao QZ, Wu HB, Qiao YS et al., 2002. Onset of Asian desertification by 22 myr ago inferred from loess deposits in China. *Nature* **416**:159-163.
- Hewitt G, 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**:907-913.
- Hewitt GM, 2004. The structure of biodiversity – insights from molecular phylogeography. *Fron Zool* **1**:4.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* **25**:1965-1978.
- Jiménez-Valverde A, Lobo JM, 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecol* **31**:361-369.

- Kearns AM, Joseph L, Toon A, Cook LG, 2014. Australia's arid-adapted butcherbirds experienced range expansions during Pleistocene glacial maxima. *Nat Commun* **5**:3994.
- Keppel G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M et al., 2012. Refugia: Identifying and understanding safe havens for biodiversity under climate change. *Glob Ecol Biogeogr* **21**:393-404.
- Kovářík F, 2019. Taxonomic reassessment of the genera *Lychas*, *Mesobuthus*, and *Olivierus*, with descriptions of four new genera (Scorpiones: Buthidae). *Euscorpius* **288**:1-27.
- Lemey P, Rambaut A, Welch JJ, Suchard MA, 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Mol Biol Evol* **27**:1877-1885.
- Li Z, Sun D, Chen F, Wang F, Zhang Y et al., 2014. Chronology and paleoenvironmental records of a drill core in the central Tengger desert of China. *Quat Sci Rev* **85**:85-98.
- Librado P, Rozas J, 2009. Dnasp v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**:1451-1452.
- Lifton N, Beel C, Hättestrand C, Kassab C, Rogozhina I et al., 2014. Constraints on the late Quaternary glacial history of the Inylchek and Sary-dzaz valleys from in situ cosmogenic ¹⁰Be and ²⁶Al, eastern Kyrgyz Tian Shan. *Quat Sci Rev* **101**:77-90.
- Liu J, Guo X, Chen D, Li J, Yue B et al., 2019. Diversification and historical demography of the rapid racerunner (*Eremias velox*) in relation to geological history and Pleistocene climatic oscillations in arid Central Asia. *Mol Phylogenet Evol* **130**:244-258.
- Liu T, Ding Z, Rutter N, 1999. Comparison of Milankovitch periods between continental loess and deep sea records over the last 2.5 Ma. *Quat Sci Rev* **18**:1205-1212.
- Lu H, Yi S, Xu Z, Zhou Y, Zeng L et al., 2013. Chinese deserts and sand fields in Last Glacial Maximum and Holocene Optimum. *Chin Sci Bull* **58**:2775-2783.
- Lü Y, Gu Z, Aldahan A, Zhang H, Possnert G et al., 2010. ¹⁰Be in quartz gravel from the Gobi desert and evolutionary history of alluvial sedimentation in the Ejina basin, Inner Mongolia, China. *Chin Sci Bull* **55**:3802-3809.
- Melville J, Hale J, Mantziou G, Ananjeva NB, Milto K et al., 2009. Historical biogeography, phylogenetic relationships and intraspecific diversity of agamid lizards in the Central Asian deserts of Kazakhstan and Uzbekistan. *Mol Phylogenet Evol* **53**:99-112.
- Meng H-H, Zhang M-L, 2013. Diversification of plant species in arid northwest China: Species-level phylogeographical history of *Lagochilus bunge* ex Benth (Lamiaceae). *Mol Phylogenet Evol* **68**:398-409.
- Middleton NJ, Sternberg T, 2013. Climate hazards in drylands: A review. *Earth Sci Rev* **126**:48-57.
- Mirshamsi O, Sari A, Elahi E, Hosseini S, 2010. Phylogenetic relationships of *Mesobuthus eupeus* (c.L. Koch, 1839) inferred from COI sequences (Scorpiones: Buthidae). *J Nat Hist* **44**:2851-2872.
- Mirshamsi O, Sari A, Elahi E, Hosseini S, 2011. *Mesobuthus eupeus* (Scorpiones: Buthidae) from Iran: A polytypic species complex. *Zootaxa* **2929**:1-21.
- Noy-Meir I, 1973. Desert ecosystems: Environment and producers. *Annu Rev Ecol Syst* **4**:25-51.
- Pfeiffer M, Chimedregzen L, Ulykpan K, 2003. Community organization and species richness of ants (Hymenoptera/Formicidae) in Mongolia along an ecological gradient from steppe to Gobi desert. *J Biogeogr* **30**:1921-1935.

- Pfeiler E, Johnson S, Richmond MP, Markow TA, 2013. Population genetics and phylogenetic relationships of beetles (Coleoptera: Histeridae and staphylinidae) from the Sonoran desert associated with rotting columnar cacti. *Mol Phylogenet Evol* **69**:491-501.
- Phillips SJ, Anderson RP, Schapire RE, 2006. Maximum entropy modeling of species geographic distributions. *Ecol Modell* **190**:231-259.
- Polis G, Yamashita T, 1991. The ecology and importance of predaceous arthropods in desert communities. In: Polis G, editor. *The Ecology of Desert Communities*. Tucson: The University of Arizona Press. p. 180-222.
- Posada D, Crandall KA, 1998. Modeltest: Testing the model of DNA substitution. *Bioinformatics* **14**:817-818.
- Provan J, Bennett KD, 2008. Phylogeographic insights into cryptic glacial refugia. *Trends Ecol Evol* **23**:564-571.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA, 2018. Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Syst Biol* **67**:901-904.
- Ramos-Onsins SE, Rozas J, 2002. Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**:2092-2100.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A et al., 2012. Mrbayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**:539-542.
- Salvi D, Schembri PJ, Sciberras A, Harris DJ, 2014. Evolutionary history of the maltese wall lizard *Podarcis filfolensis*: Insights on the ‘expansion–contraction’ model of pleistocene biogeography. *Mol Ecol* **23**:1167-1187.
- Schmitt T, 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front Zool* **4**:11.
- Shi C-M, Huang Z-S, Wang L, He L-J, Hua Y-P et al., 2007. Geographical distribution of two species of *Mesobuthus* (Scorpiones: Buthidae) in China: Insights from systematic field survey and predictive models. *J Arachnol* **35**:215-226.
- Shi C-M, Ji Y-J, Liu L, Wang L, Zhang D-X, 2013. Impact of climate changes from middle Miocene onwards on evolutionary diversification in Eurasia: Insights from the mesobuthid scorpions. *Mol Ecol* **22**:1700-1716.
- Shi C-M, Liang H-B, Altanchimeg D, Nonnaizab, Chuluunjav C et al., 2015. Climatic niche defines geographical distribution of *Mesobuthus eupeus mongolicus* (Scorpiones: Buthidae) in gobi desert. *Zool Syst* **40**:339-348.
- Shi X-J, Zhang M-L, 2015. Phylogeographical structure inferred from cpDNA sequence variation of *Zygophyllum xanthoxylon* across north-west China. *J Plant Res* **128**:269-282.
- Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ et al., 2018. Bayesian phylogenetic and phylodynamic data integration using Beast 1.10. *Virus Evol* **4**:vey016.
- Sun J, Gong Z, Tian Z, Jia Y, Windley B, 2015. Late Miocene stepwise aridification in the Asian interior and the interplay between tectonics and climate. *Palaeogeogr Palaeoclimatol Palaeoecol* **421**:48-59.
- Tajima F, 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585-595.

- Wang Y, Zhao L-M, Fang F-J, Liao J-C, Liu N-F, 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**:362-383.
- Willis KJ, Bennett KD, Walker D, Hewitt GM, 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lon B: Bio Sci* **359**:183-195.
- Worth JRP, Williamson GJ, Sakaguchi S, Nevill PG, Jordan GJ, 2014. Environmental niche modelling fails to predict Last Glacial Maximum refugia: Niche shifts, microrefugia or incorrect palaeoclimate estimates? *Glob Ecol Biogeogr* **23**:1186-1197.
- Yang S, Ding Z, 2006. Winter–spring precipitation as the principal control on predominance of C₃ plants in Central Asia over the past 1.77 myr: Evidence from $\delta^{13}\text{C}$ of loess organic matter in Tajikistan. *Palaeogeogr Palaeoclimatol Palaeoecol* **235**:330-339.
- Yin A, 2010. Cenozoic tectonic evolution of asia: A preliminary synthesis. *Tectonophysics* **488**:293-325.
- Zhang D-X, Hewitt GM, 1998. Isolation of animal cellular total DNA. In: Karp A, Isaac P, Ingram D, editors. Molecular tools for screening biodiversity: plants and animals. London: Chapman & Hall. p. 5-9.
- Zhang X, Liu G, Feng Y, Zhang D, Shi C, 2020. Genetic analysis and ecological niche modeling delimit species boundary of the Przewalski's scorpion (Scorpiones: Buthidae) in arid asian inland. *Zool Syst* **45**:81-96.

Figure legends:

Figure 1. Geographic sampling of the desert scorpion *M. mongolicus* (Mmon). The sampled sites were plotted on a digital elevation map of the studies areas. The sampled sites were also projected on the map of Köppen-Geiger climatic classification (inserted, upright). BWk: arid, desert, cold; BSk: arid, steppe, cold; Df(a-c): cold, without dry season; Dw(a-c) cold, dry winter; EF: polar, frost; ET: polar, tundra. Orange dots, the Gobi population of Mmon; orange circles, the Junggar population of Mmon; blue diamonds, *M. przewalskii*. Mmon and *M. przewalskii* are typical arid species occurring widely in arid ecosystems (BWk and BSk). Green area in the low right represents suitable distribution range of *M. martensii*.

Figure 2. The desert scorpion *M. mongolicus* (Mmon) is the most derived lineage of the *M. eupeus* species complex. (A) Maximum likelihood tree inferred from the mitochondrial DNA data (mtCOI). (B) Median-joining network of Mmon mitochondrial haplotypes. Branch lengths are proportional to the number of substitutions occurred between haplotypes and circle sizes proportional to the number of individuals sharing the haplotypes. (C) Maximum likelihood tree inferred from three concatenated nuclear genes (PK, STE and Spn2). Numbers at the nodes represents the bootstrap support value and Bayesian posterior probabilities. Me complex, *M. eupeus* species complex.

Figure 3. Spatiotemporal reconstruction of the geographic dispersal of *M. mongolicus* (Mmon) using the Bayesian phylogeographic method. Shown here are three snapshots shown that Mmon was originated in Central Asia (A), then dispersed eastward step by step (B), and finally to the Gobi regions (C). Geographic spread of genealogies across spaces was modelled with a continuous diffusion process governed by the relaxed random walk. Circles represent nodes of the maximum clade credibility phylogeny. Shaded areas represent 80% highest posterior density interval which depicted the uncertainty of phylogeographic estimates for the nodes.

Figure 4. Bayesian skyline plots derived from samples of mtCOI sequences for *M. mongolicus* (A) and *M. przewalskii* (B). The medians and their 95% highest posterior density intervals of effective population size ($N_e\tau$) were plotted against time. The x axis represents time in thousand years (kyr) before present. The y axis is the log scaled effective population size ($N_e\tau$). Population size for both species increased during the LGM about 21 kyr ago.

Figure 5. Potentially suitable distribution areas for *M. mongolicus* (Mmon) at present (A) and during the LGM (B). Ecological niche model generated under current conditions was hind projected to climatic condition at the LGM using Maxent. Orange areas represent the inferred range based on the MTSS threshold, and the yellow areas are the inferred range according to the ENT threshold. The green area in the low right represents suitable distribution range of *M. martensii*.

Table 1. Divergence time estimates generated by the relaxed molecular clock approach based on *mtCOI* data. Means and the corresponding 95% highest posterior density (HPD) intervals for times to most recent common ancestor (T_{MRCA}) were given in million years for each lineage as shown in Fig. 2.

MRCA	Mean	HPD lower	HPD upper
<i>M. mongolicus</i> (Mmon)-KZ01	3.27	2.12	4.54
<i>M. mongolicus</i> (Mmon)	1.36	0.78	1.97
Junggar-Gobi lineage	0.54	0.30	0.80
<i>M. przewalskii</i>	1.00	0.56	1.51
<i>M. przewalskii</i> - <i>M. martensii</i>	6.27	4.35	8.37
Mmon- <i>M. martensii</i>	10.56	8.33	12.70

Table 2. Summary of genetic statistics of mtCOI data for *M. mongolicus* (Mmon). Here given are sample size (N), number of haplotypes (h), hapotype diversity (Hd), nucleotide diversity (π), Tajima's *D*, Fu's *F_s*, and Ramos-Osins & Rozas' *R₂* statistics. ** $P < 0.01$, * $P < 0.05$

Populatio	N	h	Hd	π	<i>D</i>	<i>F_s</i>	<i>R₂</i>
Mmon	336	60	0.7189	0.0023	-2.3955**	-92.0339**	0.0761**
Junggar	19	7	0.7485	0.0019	-1.0758	-3.1589**	0.0909*
Gobi	317	53	0.6849	0.0020	-2.4037**	-28.8463**	0.0122**
South	255	45	0.7399	0.0023	-2.3120**	-59.5270**	0.0798**
North	62	10	0.3246	0.0006	-2.1978**	-11.9593**	0.1198**

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5