

Molecular phylogenetic and phylogeography of *Soriculus nigrescens* in southwest China

Haijun Jiang¹, Changkun Fu¹, Keyi Tang¹, Fengjun Li², Fei Xie¹, Boxin Qin¹, Dan Chen¹, Shunde Chen¹, and Shaoying Liu³

¹Sichuan Normal University

²Sichuan University

³Sichuan Academy of Forestry

March 2, 2021

Abstract

The Himalayan Shrew belongs to the genus *Soriculus*, which is a monotypic genus. We investigated the phylogeographic patterns, speciation, divergence time, and demographic history of *Soriculus nigrescens* in southwestern China. A total of 128 samples from 29 regions were analyzed for both mitochondrial and nuclear genes. Phylogenetic analyses of the concatenated mtDNA and nuDNA data revealed three highly divergent lineages within the *S. nigrescens*. One group represented subspecies *S. n. minors*, another group represented subspecies *S. n. nigrescens*, which is made up of two Clades A and B. The species delimitation analyses, based on two methods, supported the species status of the two Clades of *S. n. nigrescens*. In addition, it was found that individuals at different altitudes in Motuo were divided into two Clades. Bayesian skyline plotting analyses and ecological niche modeling also supported demographic and range expansions during the LGM for *S. n. nigrescens*. We propose that *S. n. nigrescens* appears to be composed of two putative species, and *S. n. minor* should be elevated to species status. Our study also suggested that climate change since the Miocene periods and the uplift of the QTP may have resulted in the diversification and speciation of *S. nigrescens*.

Molecular phylogenetic and phylogeography of *Soriculus nigrescens* in southwest China

Jiang haijun^{1#}, Fu changkun^{1#}, Tang keyi¹, Li fengjun², Xie fei¹, Qin boxin¹, Chen dan¹, Chen shunde^{1*}, Liu shaoying^{3*}

1College of Life Sciences, Sichuan Normal University, Chengdu 610066, China.

2College of Life Sciences, Sichuan University, Chengdu 610065, China.

3 Sichuan Academy of Forestry, Chengdu 610081, China

Authors contributed equally to this work.

*Corresponding Author: Chen shunde: csd111@126.com; Liu shaoying: Shaoyliu@163.com

Abstract

The Himalayan Shrew belongs to the genus *Soriculus*, which is a monotypic genus. We investigated the phylogeographic patterns, speciation, divergence time, and demographic history of *Soriculus nigrescens* in southwestern China. A total of 128 samples from 29 regions were analyzed for both mitochondrial and nuclear genes. Phylogenetic analyses of the concatenated mtDNA and nuDNA data revealed three highly divergent lineages within the *S. nigrescens*. One group represented subspecies *S. n. minors*, another group represented subspecies *S. n. nigrescens*, which is made up of two Clades A and B. The species delimitation

analyses, based on two methods, supported the species status of the two Clades of *S. n. nigrescens*. In addition, it was found that individuals at different altitudes in Motuo were divided into two Clades. Bayesian skyline plotting analyses and ecological niche modeling also supported demographic and range expansions during the LGM for *S. n. nigrescens*. We propose that *S. n. nigrescens* appears to be composed of two putative species, and *S. n. minor* should be elevated to species status. Our study also suggested that climate change since the Miocene periods and the uplift of the QTP may have resulted in the diversification and speciation of *S. nigrescens*.

Keywords: The Himalayan Shrew; Population demography; Species delimitation; Population expansion

Introduction

It seems that the speciation driving patterns, genetic variation and population structure of many species are all affected by paleoclimate fluctuations and paleogeological events (Avise, 2009; He and Jiang, 2014; Feng et al., 2016; Ye et al., 2016a, 2016b). Both current and historical environmental factors, such as topography, climate and habitat, were strongly affected evolution of small mammals (Kent et al., 2014). The shrews have limited dispersal abilities due to their small body, high metabolism, and low vagility (Churchfield, 1990; Churchfield et al., 2012). As a consequence, the evolution of terrestrial shrews were strongly affected by Climatic and Topographic situations (Vega et al., 2010; Wierzbicki et al., 2011; Willows-Munro and Matthee, 2011). Shrews could provide excellent models for exploring the genetic consequences of climatic changes and topographic isolation (Chen et al., 2015).

The distribution pattern of biodiversity and its causes are one of the core contents of biogeography research, which is of great significance for understanding the evolutionary laws of biodiversity, the process of species formation, and formulating the protection and management of species diversity (Chen et al., 2015). There are a total of 25 ecological hotspots in the world, all of which have high biodiversity (Myers et al., 2000). Southwest China is a hotspot of biodiversity in East Asia due to the high level of endemism and extensive biodiversity (He and Jiang, 2014). Geological and climatic events affect the morphology, spatial structure and genetic structure of animals, especially for thermostatic vertebrates (Wu et al., 2013). The uplift of the Qinghai-Tibet Plateau (QTP) and the Quaternary glacial period, orogeny, the formation of the Asian monsoon and the violent Quaternary climate oscillations have profoundly affected the distribution and genetic structure of existing organisms (An et al., 2001; He and Jiang, 2014; Lei et al., 2014; Qu et al., 2015; Ye et al., 2016a,b). During the ice age, although the ice sheet was not completely covered, it had a huge impact on species.

Research by Reumer (1987) indicated that temperature and humidity are the key factors affecting the distribution, number and species of shrews. The Himalayan Shrew (*Soriculus nigrescens*) Gray, 1842 belongs to the Soricinae, which is distributed in southwest China, Nepal, India and Bhutan (Smith and Xie, 2010). And it is a monotypic genus. The classification of *Soriculus nigrescens* was confusing before. Based on morphological data, Motokawa (2003) divided *Soriculus nigrescens* into two subspecies, *S. n. nigrescens* (Gray, 1842) and *S. n. minor* Dobson, 1890 (Motokawa, 2003). *Soriculus nigrescens minor* was considered as senior synonym of *S. radulus* Thomas, 1922, and *S. n. nigrescens* was regarded as senior synonyms of *S. aterrimus*, *S. sikimensis*, *S. n. pahari*, *S. n. caurinus*, and *S. n. centralis* (Motokawa, 2003). But currently there is a lack of molecular phylogenetic research data on *Soriculus nigrescens*.

In this study, we analyzed the genetic structure of the population, calculated the time of differentiation and the historical dynamics of the population, reshaped the niche model of *Soriculus nigrescens*, and studied the relationship between its lineages and system evolution, in order to explore the genetic diversity and the mechanism of species diversity formation in the Himalayas in China and provide data for differentiation mechanism of small mammals in the Himalayas.

Materials and methods

Sampling, sequencing and alignment

From 2009 to 2019, a total of 124 samples of the *S. n. nigrescens* were collected from 28 regions in Tibet

of China. Three samples of the *S. nigrescens* from Yunnan were downloaded from the Genbank. Additional sequence data of Soricidae were downloaded from the GenBank and included in the analysis (Table S1). Detailed specimen and locality data are presented in Fig 1A and Table S2.

Total genomic DNA was extracted from muscles or livers using a chloroform procedure (Sambrook and Russell, 2006). The mitochondrial Cyt-B gene, nuclear genes APOB, BRCA-1 and RAG-2 were amplified by highly conserved primers, detailed primer data are presented in Appendix Table S3. The same PCR primers and primer walking were used for sequencing, detailed PCR amplification and sequencing methods followed Chen et al., (2015). The Cyt-B, APOB, BRCA-1 and RAG-2 sequences were aligned and examined by eye in MEGA v5 (Tamura, et al., 2011).

Genetic diversity and structure

In this study, DnaSP 5.10 was used to calculate the haplotype number (H), haplotype polymorphism, haplotype diversity (HD), and nucleotide diversity (π) of the sample, to analyze the polymorphic data of DNA (Librado and Rozas, 2009). Estimated genetic divergence among the major mtDNA lineages (see below) through the paired Kimura two-parameter (K2P) model implemented in MAGA v5.0 (Tamura et al., 2011). Each nuclear gene calls the PHASE typing algorithm (Stephens and Donnelly, 2003) in DnaSP v5.10 (Librado and Rozas, 2009), and uses the MP method to remove redundant connections. Haplotype median-joining networks (MJN) of genealogical relationships for nuDNA were also constructed with the software Network v4.6 (Bandelt et al., 1999).

Phylogenetic analyses

We used BEAST v1.6.1 for Bayesian phylogenetic reconstructions (Drummond and Rambaut, 2007). Bayesian gene trees were reconstructed independently for concatenated Cyt-B and each nuDNA gene. The jModeltest v2 (Darriba et al., 2012) was used to determine the model of evolution for mtDNA and nuDNA genes by Bayesian Information Criterion (BIC)(Table S4). The prior value of the parameter uses the default value, the starting tree uses a random spanning tree, and each data set is run at least 4 times. Each operation is 100 million generations, and every 5,000 generations are sampled. We used Tracer v1.5 to analyze the log file in the calculation result to check the convergence of the calculation. When the posterior value of the data set is greater than 200 (ESS[?]200), the analysis is considered stable and the result is credible (Drummond and Rambaut, 2007). After that, according to the posterior distribution, Tree Annotator is used to determine the burn-in. In addition, when using Figtree v1.4.3 to observe Bayesian trees, only when the DNA sequences of certain groups are clustered on nodes with a posterior probability value greater than 0.95 (PP > 0.95), it is proved that the group is highly supported (Huelsenbeck and Rannala, 2004).

Species delimitation

First, we constructed the species tree by using the *BEAST model (Heled and Drummond, 2008). The species tree was reconstructed for *S. nigrescens* based on a coalescent-based method implemented in *BEAST. *Chodsigoa hypsibius* was selected as outgroup. Only individuals with both mtDNA and nuDNA data were included in the species delimitation analysis. In order to meet the construction requirements of the *BEAST model, the model we adopted is the same as the phylogenetic analysis. *BEAST program setting: The main parameter setting is consistent with the analysis of the divergence time, the running generation is modified to 100 million generations, and every 5,000 generations are sampled.

Second, we used Bayesian species-delimitation method to delineate species boundaries (Yang and Rannala, 2014). We tested the validity of our assignment of 2 putative species (Clade A and Clade B) in BPP v3.1. Only dataset 1 (nuDNA) and dataset 2 (mtDNA + nuDNA) were included in the analysis. Two alternative rjMCMC algorithms (algorithms 0 and 1) implemented in BPP. We used algorithm 0 with fine tuning parameters for subsequent analysis as the original analysis showed that algorithms 0 and 1 produce similar results. We used Gamma-distributed priors (G) to specify the ancestral population size (θ) and root age (τ). The three combinations were modeled to allow a range of speciation histories: G (2, 2000) for θ and G (2, 2000) for τ , G (1, 10) for θ and G (2, 2000) for τ and G (1, 10) for θ and G (1, 10) for τ .

Each rjMCMC was run for 100,000 generations and sampled each 100 generations after discarding 10,000 generations as pre-burn-in.

Divergence time estimates

Cyt-B, APOB and BRCA-1 were used to calculate the divergence time of *S. nigrescens* population through the BEAST v1.6.1 (Drummond, et al., 2012). Following He et al., (2018), we set the mean to 36 Ma, with a standard deviation of 0.135, so that the median of the prior was 35.7 Ma and the 95% CI was 28.6 - 44.5 Ma. BEAST analyses used unlinked substitute model, linked clock models, linked tree, a random starting tree, a birth-death process tree prior, a relaxed lognormal clock model, and the programs default prior distributions of model parameters. Each analysis was run for 100 million generations and sampled every 5,000 generations. Convergence was accessed using Tracer v1.4 to confirm the effective sample sizes (ESS) of all parameters were higher than 200 (Rambaut and Drummond, 2007). The first 10% of the trees were discarded as the burn-in.

Population demography

We used Bayesian skyline plots (BSP) implemented in BEAST v1.6.1 (Drummond et al., 2012) to examine the historical demographics of *S. n. nigrescens* with the Cyt-B and nuDNA data. The HKY + G and GTR + I + G nucleotide substitution models selected by jModeltest 0.1 were used. The BSP analysis was executed iterations with 100 million and sampled at every 5000th tree, with the first 10% discarded as the burn-in. Convergence was accessed using Tracer v1.4 to confirm the effective sample sizes (ESS) of all parameters were higher than 200 (Rambaut and Drummond, 2007).

Ecological niche analysis

We used ecological niche modeling (ENM) to predict the distribution of *S. n. nigrescens* (Clades A, B in Fig. 1B) at the present time and during the last glacial maximum (LGM). The maximum entropy algorithm implemented in MAXENT v3.3.3 (Phillips et al., 2006; Phillips and Dudík, 2008) was used to model the ecological niches. A total of 28 distribution localities were collected. We used three environmental layers in different periods: Last Interglacial (LIG), Last Glacial Maximum (LGM) and Present. The environment layer comes from WorldClim (Hijmans et al., 2005) (website <http://www.Worldclim.Org>). The main climatic elements considered and reflected are monthly precipitation and average precipitation, minimum and maximum temperature, seasonal variation and limit level. The Maxent analysis was used with 80% of the species records for training and 20% for testing the model, 10 subsample replicates and a maximum of 2000 iterations. The area under the curve (AUC) of the receiver operating characteristic (ROC) plot was used for model evaluation.

RUSULT

Sequence characteristics and genetic diversity

The 127 individuals used in this study were from 28 sampling sites, and 3,093 bp of gene sequence was obtained after sequencing, including 1,953 bp of nuclear gene sequence (APOB [513 bp], BRCA-1 [765 bp], RAG-2 [675 bp]) and 1,140 bp mitochondrial sequence (Cyt-B [1,140 bp]). All coding genes fragments were tested in MAGA v5 (Tamura et al., 2011), and the results showed that there was no premature termination of coding genes. The mitochondrial gene Cyt-B has the most mutation sites compared with nuclear genes, and mitochondrial genes have higher genetic polymorphisms. For all the *Soriculus nigrescens* samples, the haplotype diversity (h) and nucleotide diversity (π) of mitochondria are relatively high ($h = 0.97141$, $\pi = 0.07291$). Among all amplified genes, mitochondrial Cyt-B had the highest haplotype, and 51 haplotypes were identified in Cyt-B. Among nuclear genes, RAG-2 has the highest haplotype number, and 25 haplotypes were identified in RAG-2. The nucleotide diversity and haplotype diversity of mitochondrial genes of Clade B are significantly higher than those of Clade A (Table S5). The average distances between species were 7.0% (Clade A and Clade B), 12.4% (Clade A and Clade YN), and 11.3% (Clade B and Clade YN) (Table 1).

Phylogenetic relationship and population genetic structure

The phylogenetic tree results of the mitochondrial data set, nuclear gene data set, and the combined data set showed that *S. n. nigrescens* can be divided into three clearly Clades (Clade A, Clade B and Clade YN), with higher posterior probability (Fig 2). But the phylogenetic status of Clade YN is not completely consistent in the three data sets. But Clade A and Clade B are always the closest relatives. Clade A is from Yadong County, Dingri County, Nielamu County, and Motuo County in Tibet. Clade B are from Bomi County, Milin County, Gongbujiangda County, Bayi District and Motuo County in Tibet. In Motuo County, Clade A and Clade B exist at the same time, Clade A is in the low-altitude area, and Clade B is in the high-altitude area. The MJN for the nuDNA data showed a similar structure to the phylogenetic tree, Clade A, Clade B and Clade YN have no crossover phenomenon and can be completely divided into three Clades.

We used the nuclear genes APOB, BRCA-1, and RAG-2 of Clade A and Clade B to analyze the mediation network diagram to obtain 3 MP trees. The results showed that the MP tree results of all nuclear genes of *S. n. nigrescens* are consistent with the previous mitochondrial phylogenetic tree results, and Clade A and Clade B are clearly divided into two clades (Fig S1). In all nuclear gene mediation network diagrams, Clade A and Clade B have no crossover phenomenon, and can be completely divided into two branches, which is consistent with the results of the previous nuclear gene phylogeny diagram.

Species delimitation

The *BEAST species-tree and BPP analyses gave identical results. The results of the *BEAST species-tree delimitation analysis of individuals with both the nuDNA (APOB, BRCA-1 and RAG-2) and Cyt-B data are shown in Fig 3A and Table S6. The results of the *BEAST species tree supported the *S. nigrescens* as three species, and *S. n. nigrescens* divided into two species. When using either dataset 1 (nuDNA) and dataset 2 (mtDNA + nuDNA) of *S. n. nigrescens*, BPP consistently supported the validity of 2 hypothetical species with high posterior probabilities (PP [?] 0.96)(Table S6).

Divergence time estimation

Divergence dating estimated that *S. n. nigrescens* diverged with *S. n. minor* 11.11 MYA during the early late Miocene. The divergence time of *S. n. nigrescens* further differentiated into two clades of Clade A and Clade B was at 5.69 MYA during the late Miocene (PP [?] 0.95). The results showed that it was affected by QTP change. The first uplift of QTP promoted the differentiation of *S. n. nigrescens* and *S. n. minor*. And the second uplift of QTP promoted the differentiation of Clade A and Clade B. It is consistent with the uplift of QTP in the Late Miocene. Since then, Clade A has differentiated several times in the Pliocene epoch (PP [?] 0.95); Clade A has an earlier differentiation time than Clade B, which is also in the Pliocene epoch (PP [?] 0.95)(Fig 3B).

Historical demography

In the combined analysis of Clade A and Clade B, the results of the BSP combined with nuclear gene and mitochondrial gene analysis show that the curve gradually rises from the present to the LGM (26.5 - 19.0 ka), indicating that the population expansion of *S. n. nigrescens* has occurred (Fig 4A). In the analysis of mismatch distribution, the mitochondrial gene had multiple jagged peaks, and all of nuclear gene mismatch distribution showed unimodal distributions (Fig 4B).

Ecological niche modeling

In this study, the ENM of *S. n. nigrescens* in three different climatic environments (current, LGM and LIG) have achieved good prediction results (Fig 1B). The results of divergence time show that the history of differentiation between the two branches of Clade A and Clade B can be traced back to the Late Miocene. The results of ENM show that the population size of each branch of the *S. n. nigrescens* migrated and changed during the severe fluctuations of the last glacial maximum and the last interglacial period. Since the LIG, the distribution range of *S. n. nigrescens* has gradually expanded. Compared with the BSP results, the ENM results show that the population expansion time may be earlier.

Discussion

Biogeographic inferences

First, our study supports allopatric occurrence of species within the *S. n. nigrescens*. In this study, these species have obvious systematic geographic characteristics, and their distribution ranges do not overlap. Clade A is mainly distributed in the altitude areas of Yadong County, Nielamu County, Dingri County and Motuo County in Tibet. In the Motuo area, the collected specimens are deciduous broad-leaved forest areas below 1200m altitude. And there are also specimens of the *S. n. nigrescens* in the residential areas of Motuo County. The specimens of Clade B were collected from Milin, Bomi, Gongbujiangda, Bayi and Motuo. And these samples from the Motuo area came from the high-altitude area of Motuo, which is bordering the Milin. Clade YN was collected from Pianma, Yunnan Province (He and Jiang, 2014). The divergence of these lineages may be to the uplift of the QTP and the accompanying climate fluctuations. The QTP experienced four phases (12 - 14 Ma (Turner et al., 1993; Coleman and Hodges, 1995; Dettman et al., 2003; Lu et al., 2004; Sun et al., 2005), 6 - 8 Ma (Molnar et al., 1993; Molnar, 2005; Song et al., 2007), 3.6 - 4 Ma (Li and Fang, 1999; Song et al., 2007; Nie et al., 2008) and 1.2 - 0.6 Ma (Li and Fang, 1999). Each phase of a geological event can lead to habitat fragmentation, which may prevent spread and promote vicariant speciation in many taxa (He et al., 2010; Lei et al., 2014). The first uplift of QTP promoted the differentiation of *S. n. nigrescens* and *S. n. minor*. And the second uplift of QTP promoted the differentiation of Clade A and Clade B. The rapid rise of QTP and the evolution of the Asian monsoon have reshaped the landscape of East Asia.

Second, we found that climate change in the Miocene affected the genetic pattern and population history of *S. n. nigrescens*. Climate change and cyclical Quaternary glaciations have been revealed as the major factors affecting the demographic history of many extant species in Europe and North America (Hewitt, 1996, 2004; Ursenbacher et al., 2008; Fijarczyk et al., 2011). BSP demographic analyses indicated population expansion for *S. n. nigrescens*. The atypical star-like radiation from the MJN analysis implied the presence of a bottleneck for the two clades in the past. Our ENMs analyses revealed the predicted range has changed between the LGM and the present for *S. n. nigrescens* (Fig. 1B). Furthermore, the BSP analyses showed that the expansion events began from approximately 0.02 MYA, suggesting the role of Pleistocene climatic changes in driving the demographic expansion of the *S. n. nigrescens*.

Nuclear and mtDNA genetic structure

The analyses based on the Cyt-B and nuDNA data revealed a pattern which is consistent with the distribution for the *S. n. nigrescens*. The results of Cyt-B and nuDNA both support the existence of *S. n. nigrescens* as two clades.

The inconsistency between mitochondrial and nuclear data may be caused by many factors. The incomplete lineages ordering of ancestral polymorphism and the different evolution rate of different markers may cause inconsistencies between mitochondrial and nuclear gene data (Toews and Brelsford, 2012; Lin et al., 2014; Ye et al., 2016b). Incomplete lineages division is often used to explain the inconsistency between mitochondrial and nuclear results. If incomplete lineages sorting results in the observed pattern of mitochondrial DNA, it can be expected that the divergence pattern of nuclear DNA will be shallow, as a relatively large effective population size will result in a relatively long sorting time (Ye et al., 2016b). We found less nuclear genetic divergence compared to the mtDNA, and this is consistent with incomplete lineage sorting. The nuDNA appeared to be conservative and have lower evolutionary rate than mtDNA, which underestimated the genetic structure of *S. n. nigrescens*. Moreover, the rate of evolution may also have a strong impact on the ability to detect differentiation. The slow rate of evolution may not produce enough polymorphisms, and the existing genetic structure cannot be well recognized, and the excessively high rate of evolution may even mask the signals of high differentiation (Hedrick, 1999; Brito, 2007). The nuDNA appeared to be conservative and have lower evolutionary rate than mtDNA, which underestimated the genetic structure of *S. n. nigrescens*. Thus, the different evolution rates of mtDNA and nuDNA markers may lead to the inconsistency of nuclear mitochondria revealed in this study.

Taxonomic consideration

Traditional taxonomy is based on the morphological characteristics of species for identification and description, which may underestimate the species diversity due to the phenotypic plasticity and local morphological adaptations to the similar environmental habitats (Wan et al., 2013). There are cryptic species of mammals that would likely not be recognized based solely on classical studies of morphology of voucher specimens housed in museums (Kaliontzopoulou et al., 2010; Barley et al., 2013). Motokawa (2003) divided *S. nigrescens* into two subspecies, *S. n. minor* and *S. n. nigrescens* based on morphological judgment. In this study, we further divide *S. n. nigrescens* into two clades, Clade A and Clade B. Each of them formed a well-supported clade on both the Cyt-B and nuclear gene trees. And species delimitation based on the method BPP confirmed their species status with high supports. In addition, by extensive sampling, our mtDNA genealogy and MJN results all supported the existence of two highly supported genetic clades (Clade A, B) for *S. n. nigrescens*. The pairwise distance among the three clades (0.070, 0.113, 0.124) is obviously higher than the divergence threshold for identifying mammal cryptic species. Since their non-overlapping geographic distributions, it suggested the extensive lineage diversification in the *S. n. nigrescens*. The BPP species delimitation analysis indicated the two clades were candidate species. Furthermore, Clade YN always clustered separately in molecular analysis, and it does not overlap with other two clades. Therefore, the species delimitation results in this study reveal that there are two species in the *S. n. nigrescens*, and *S. n. minor* should be elevated to species status.

The type specimen of *S. n. nigrescens* was collected from India, West Bengal, Darjeeling, and kept in the Natural History Museum (London). Among the collection areas in this study, Yadong County is the area where is closest to the type specimen production area of *S. n. nigrescens*. The samples collected in Yadong County were identified as Clade A. In addition, the samples collected from Nielamu County were also identified as Clade A. In this study, we found that samples collected from central Nepal were clustered with Clade A. Therefore, we hypothesized that the specimens of Clade A and the type specimens were of the same species.

Conclusions

In the present study, we obtained molecular sequences of *S. nigrescens* in Tibet and Yunnan province. We analyzed the mtDNA and nuDNA data to study the diversification, phylogeographic structure, and evolutionary history of *S. nigrescens*. We detected three highly supported Clades within the *S. nigrescens* based on mtDNA and nuDNA genealogy analysis. We propose that *S. n. nigrescens* appears to be composed of two putative species (Clade A and Clade B), and *S. n. minor* should be elevated to species status. The urgent morphological description needs to further verify our hypothesis. Our study also suggested that climate change since the Miocene periods and the uplift of the QTP may have resulted in the diversification and speciation of *S. nigrescens*.

Figures and tables legends

Fig 1A: Map of Southern China showing the sampling localities of the *S. nigrescens* included in this study; Fig 1B Ecological niche model of *S. n. nigrescens* at LIG (c), LGM (b) and Present (a) represent.

Fig 2: Results of Bayesian phylogenetic analyses of nuclear, mtDNA and mtDNA+nuDNA and gene data sets (A mtDNA, B nuDNA, C mtDNA + nuDNA).

Fig 3: Results of species delimitation species trees using the BEAST model.

Fig 4: Bayesian phylogenetic analyses and divergence times for *S. nigrescens* based on combined genes.

Fig 5A: Extended Bayesian skyline plot of *S. n. nigrescens* based on mtDNA+nuDNA; Fig 5B: Expected (solid line) and observed mismatch distribution of *S. n. nigrescens*.

Table 1: Mean Cyt-B distance between *S. nigrescens* species calculated using the Kimura 2-parameter model.

Appendix Figure and tables legends

Fig S1: Median-joining networks of the three nuDNA genes.

Table S1: Sequence downloaded for molecular analysis.

Table S2: Sample site, numbers of specimen and gene amplification of *S. nigrescens*.

Table S3: Primers used for PCR and sequencing.

Table S4: The best molecular evolution model used in phylogenetic reconstructions.

Table S5: Sequence statistics of samples used in present study. N, H, HP and π means the number of sequences, haplotypes, haplotype diversity, nucleotide diversity, respectively.

Table S6: Posterior probabilities for the 2 putative species using different priors for model parameters.

Ethics

No permits or ethics committee approvals were required for this research. We received permission from all landowners and managers to conduct research.

Competing interests

We declare we have no competing interests.

Acknowledgements

We are very grateful for the funding from the National Natural Science Foundation of China. We are very grateful to Liao Rui from Sichuan Academy of Forestry for providing sample for this study.

Author contributions

Jiang haijun: Data curation (equal); Formal analysis (lead); Investigation (equal); Validation (equal); Visualization (lead); Writing-original draft (lead). Fu changkun: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (supporting); Writing-review & editing (lead). Tang keyi: Data curation (equal); Formal analysis (equal); Investigation (equal); Writing-review & editing (equal). Li fengjun: Data curation (equal); Investigation (equal); Writing-review & editing (equal). Xie fei: Data curation (equal); Investigation (equal); Writing-review & editing (equal). Qin boxin: Data curation (equal); Investigation (equal); Writing-review & editing (equal). Chen dan: Data curation (equal); Investigation (equal); Writing-review & editing (equal). Liu shaoying: Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (lead); Writing-review & editing (equal). Chen shunde: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (lead); Resources (lead); Supervision (equal); Validation (equal); Writing-original draft (equal).

Funding

This research was funded by the National Natural Science Foundation of China (31670388; 31970399; 32070424).

Data Accessibility

DNA sequences: Genbank accessions MW669978 - MW670101; MW670102 - MW670207; MW670208 - MW670322; MW670323 - MW670437.

References

- An ZS, Kutzbach JE, Prell WL, Porter SC. 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. *Nature* 411, 62-66. (doi.org/10.1038/35075035)
- Avice JC. 2009. Phylogeography: retrospect and prospect. *J. Biogeogr.* 36, 3-15. (doi: 10.1111/j.1365-2699.2008.02032.x)

- Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology & Evolution*. 16, 37-48. (doi: 10.1093/oxfordjournals.molbev.a026036)
- Barley A J, White J, Diesmos AC, Brown RM. 2013. The Challenge of species delimitation at the extremes: diversification without morphological change in philippine sun skinks. *Evolution* 67, 3556-3572. (doi: 10.1111/evo.12219)
- Brito PH. 2007. Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Mol Ecol*. 16, 3423-3437. (doi: 10.1111/j.1365-294X.2007.03401.x)
- Chen S, Sun Z, Kai H, Jiang X, Yang L, Koju NP, Zhang X, Tu B, Ban Z, Liu S. 2015. Molecular phylogenetics and phylogeographic structure of *Sorex bedfordiae* based on mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics Evolution* 84:245-253. (doi: 10.1016/j.ympev.2014.12.016)
- Chen Shunde, Qing Jiao, Liu Zhu, Liu Yang, Tang Mingkun, Murphy Robert, Pu, Yingting, Wang Xuming, Keji Guo, Jiang, Xuelong, Liu, Shaoying. 2019. Multilocus phylogeny and cryptic diversity of white-toothed shrews (Mammalia, Eulipotyphla, *Crocidura*) in China. (doi: 10.21203/rs.2.14115/v1)
- Churchfield S. (1990). *The Natural History of Shrews*. Comstock Publishing Associates, Ithaca, New York.
- Churchfield S, Rychlik L, Taylor JRE. (2012). Food resources and foraging habits of the common shrew, *Sorex araneus*: does winter food shortage explain Dehnel's phenomenon? *Oikos* 121, 1593-1602. (doi: 10.1111/j.1600-0706.2011.20462.x)
- Coleman M, Hodges K. 1995. Evidence for Tibetan plateau uplift before 14 Myr ago from a new minimum age for east-west extension. *Nature*, 374, 49-52. (doi: 10.1038/374049a0)
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. JModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772-772. (doi: 10.1038/nmeth.2109)
- Dettman DL, Fang X, Garzione CN, Li J. 2003. Uplift-driven climate change at 12 Ma: a long δ 18 O record from the NE margin of the Tibetan plateau. *Earth and Planetary Science Letters*, 214, 267-277. (doi: 10.1016/S0012-821X(03)00383-2)
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology*, 7(1):214. (doi: 10.1186/1471-2148-7-214)
- Drummond Alexei J, Suchard Marc A, Xie Dong, Rambaut Andrew. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution*, 29(8). (doi: 10.1093/molbev/mss075)
- Feng B, Zhao Q, Xu JP, Qin J, Yang ZL. 2016. Drainage isolation and climate change-driven population expansion shape the genetic structures of *Tuber indicum* complex in the Hengduan Mountains region. *Rep.* 6, 21811. (doi: 10.1038/srep21811)
- Fijarczyk A, Nadachowska K, Hofman S, Litvinchuk SN, Babik W, Stuglik M, Gollmann G, Choleva L, Cogalniceanu D, Vukov T, Dzukic G, Szymura, JM. 2011. Nuclear and mitochondrial phylogeography of the European fire-bellied toads *Bombina bombina* and *Bombina variegata* supports their independent histories. *Molecular Ecology*, 20(16). (doi: 10.1111/j.1365-294X.2011.05175.x)
- He K, Hu NQ, Chen X, Li JT, Jiang XL. 2016. Interglacial refugia preserved high genetic diversity of the Chinese mole shrew in the mountains of southwest China. *Heredity* 116(1):23. (doi: 10.1038/hdy.2015.62)
- He K, Li YJ, Brandley MC, Lin LK, Jiang XL. 2010. A multi-locus phylogeny of Nectogalini shrews and influences of the paleoclimate on speciation and evolution. *Molecular Phylogenetics & Evolution*, 56(2):734-746. (doi: 10.1016/j.ympev.2010.03.039).
- He K, Jiang, XL. 2014. Sky islands of southwest China. I: an overview of phylogeographic patterns. *Chinese Sci. Bull.* 59, 585-597. (doi: 10.1007/s11434-013-0089-1)

- Hedrick PW. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53, 313-318. (doi: 10.2307/2640768)
- Heled J, Drummond AJ. 2008. Bayesian inference of population size history from multiple loci. *Bmc Evolutionary Biology* 8:289-289. (doi: 10.1186/1471-2148-8-289)
- Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*. (doi: 10.1111/j.1095-8312.1996.tb01434.x)
- Hewitt GM. 2004. The Evolutionary Legacy of the Ice Ages || Genetic Consequences of Climatic Oscillations in the Quaternary. *Philosophical Transactions Biological Sciences*, 359(1442):183-195. (doi: 10.1098/rstb.2003.1388)
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005 Worldclim: High resolution interpolated surfaces for global land areas. *International Journal of Climatology*. 25:1965-1978. (doi: 10.1002/joc.1276)
- Huelsenbeck, John, Rannala, Bruce. 2004. Frequentist Properties of Bayesian Posterior Probabilities of Phylogenetic Trees Under Simple and Complex Substitution Models. *Systematic Biology*, 53(6):904-913. (doi: 10.1080/10635150490522629)
- He Kai, Chen Xing, Chen Peng, He Shui-Wang, Cheng Feng, Jiang Xue-Long. 2018. A new genus of Asiatic short-tailed shrew (Soricidae, Eulipotyphla) based on molecular and morphological comparisons. *Zoological Research* (5), 321-334. (doi: 10.24272/j.issn.2095-8137.2018.058)
- Kaliontzopoulou A, Carretero M A, Neftalí Sillero. 2010. Geographic patterns of morphological variation in the lizard *Podarcis carbonelli*, a species with fragmented distribution. *Herpetological Journal*, volume 20(20):41-50.
- Kent R, Bar-Massada A, Carmel Y. (2014). Bird and mammal species composition in distinct geographic regions and their relationships with environmental factors across multiple spatial scales. *Ecol. Evol.* 4, 1963–1971. (doi: 10.1002/ece3.1072)
- 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. *Current Zoology*. 60, 149-161. (doi: 10.1093/czoolo/60.2.149)
- Li J, Fang X. 1999. Uplift of the Tibetan Plateau and environmental changes. *Chinese Science Bulletin*, 44, 2117-2124. (doi: 10.1007/BF03182692)
- Librado, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451-1452. (doi: 10.1093/bioinformatics/btp187)
- Lin AQ, Csorba G, Li LF, Jiang TL, Lu GJ, Thong VD, Soisook P, Sun KP, Feng J. 2014. Phylogeography of *Hipposideros armiger* (Chiroptera: Hipposideridae) in the Oriental Region: the contribution of multiple Pleistocene glacial refugia and intrinsic factors to contemporary population genetic structure. *Journal of Biogeography*, 2014, 41, 317-327. (doi: 10.1111/jbi.12163)
- Lu HY, Wang XY, An ZS, Miao XD, Zhu, RX, Ma HZ, Li Z, Tan HB, Wang XY. 2004. Geomorphologic evidence of phased uplift of the northeastern Qinghai-Tibet Plateau since 14 million years ago. *Science China Earth Sciences*, 47, 822-833. (doi: 10.1360/03yd0315)
- Molnar P. 2005. Mio-Pliocene growth of the Tibetan Plateau and evolution of East Asian climate. *Palaeontologia Electronica*, 8, 1-23. (doi:10.1016/j.palaeo.2009.11.002)
- Molnar P, England P, Martinod J. 1993. Mantle dynamics, uplift of the Tibetan Plateau, and the Indian monsoon. *Reviews of Geophysics*, 31, 357-396. (doi: 10.1029/93RG02030)
- Motokawa, M. 2003. *Soriculus minor* Dobson, 1890, senior synonym of *S. radulus* Thomas, 1922 (Insectivora, Soricidae). *Mammalian Biology*, 68(3), 178-180. (doi: 10.1078/1616-5047-00080)

- Myers N, Mittermeier RA, Mittermeier CG, Bonseca GA, Da, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403:853-858. (doi: 10.1038/35002501)
- Nie J, King JW, Fang X. 2008. Tibetan uplift intensified the 400 ky signal in paleoclimate records at 4 Ma. *Geological Society of America Bulletin*, 120, 1338-1344. (doi: 10.1130/B26349.1)
- Ohdachi S, Masuda R, Abe H, Dokuchaev NE. 1997. Biogeographical history of northeastern asiatic soricine shrews (insectivora, mammalia). *Researches on Population Ecology*, 39(2), 157-162. (doi: 10.1007/BF02765261)
- Ohdachi SD, Hasegawa M, Iwasa MA, Vogel P, Oshida T, Lin LK. 2010. Molecular phylogenetics of soricid shrews (mammalia) based on mitochondrial cytochrome b gene sequences: with special reference to the soricinae. *Proceedings of the Zoological Society of London*, 270(1), 177-191. (doi: 10.1111/j.1469-7998.2006.00125.x)
- Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *ECOLOGICAL MODELLING*. (doi: 10.1016/j.ecolmodel.2005.03.026)
- Phillips SJ, Miroslav Dudik. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2):161-175. (doi: 10.1111/j.0906-7590.2008.5203.x)
- 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. *Biogeogr.* 42, 179-192. (doi: 10.1111/jbi.12407)
- Reumer JWF. 1987. Redefinition of the Soricidae and the Heterosoricidae (Insectivora, Mammalia), with the description of the Crocidosoricinae, a new subfamily of Soricidae. *Revue de Paleobiologie*, 6(2):189-192.
- Sambrook J, Russell DW. 2006. *Molecular Cloning*, 4th ed. Cold Spring Harbor Laboratory Press, New York.
- Smith AT and Xie Yan . 2010. *A Guide to the Mammals of China*. Princeton University Press .
- Song Y, Fang X, Torii M, Ishikawa N, Li J, An Z. 2007. Late Neogene rock magnetic record of climatic variation from Chinese eolian sediments related to uplift of the Tibetan Plateau. *Journal of Asian Earth Sciences*, 30, 324-332. (doi: 10.1016/j.jseaes.2006.10.004)
- Stephens M, Donnelly P. 2003. A Comparison of Bayesian Methods for Haplotype Reconstruction from Population Genotype Data. *American Journal of Human Genetics* 73:1162-1169. (doi: 10.1086/379378)
- Sun J, Zhu R, An Z. 2005. Tectonic uplift in the northern Tibetan Plateau since 13.7 Ma ago inferred from molasse deposits along the Altyn Tagh Fault. *Earth and Planetary Science Letters*, 235, 641-653. (doi: 10.1016/j.epsl.2005.04.034)
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MAGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* 28:2731-2739. (doi: 10.1093/molbev/msr121)
- Wan Tao, He Kai, Jiang Xue-Long. 2013. Multilocus phylogeny and cryptic diversity in Asian shrew-like moles (*Uropsilus*, Talpidae): implications for taxonomy and conservation. *BMC evolutionary biology*, 13, 232. (doi: 10.1186/1471-2148-13-232)
- Toews DPL , Brelsford A . 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21(16). (doi: 10.1111/j.1365-294X.2012.05664.x)
- Turner S, Hawkesworth C, Liu J, Rogers N, Kelley S, van Calsteren P. 1993. Timing of Tibetan uplift constrained by analysis of volcanic rocks. *Nature*, 364, 50-54. (doi: 10.1038/364050a0)
- Ursenbacher S, Schweiger S, L Tomović, J Crnobrnja-Isailović, Fumagalli L, Mayer W. 2008. Molecular phylogeography of the nose-horned viper (*Vipera ammodytes*, Linnaeus (1758)): evidence for high genetic

diversity and multiple refugia in the Balkan peninsula. *Molecular Phylogenetics & Evolution*, 46(3):1116-1128. (doi: 10.1016/j.ympev.2007.11.002)

Vega, R., Fløjgaard, C., Lira-Noriega, A., Nakazawa, Y., Svenning, J.C., Searle, J.B., 2010. Northern glacial refugia for the pygmy shrew *Sorex minutus* in Europe revealed by phylogeographic analyses and species distribution modelling. *Ecography* 33, 260–271. (doi: 10.1111/j.1600-0587.2010.06287.x)

Wierzbicki H, Moska M, Strzala T, Macierzyn´ska A. 2011. Do aquatic barriers reduce male-mediated gene flow in a hybrid zone of the common shrew (*Sorex araneus*)? *Hereditas* 148, 114–117. (doi: 10.1111/j.1601-5223.2010.02224.x)

Willows-Munro S, Matthee CA. 2011. Linking lineage diversification to climate and habitat heterogeneity: phylogeography of the southern African shrew, *Myosorex varius*. *J. Biogeogr.* 38, 1976–1991. (doi: 10.1111/j.1365-2699.2011.02543.x)

Wu SJ, Luo J, Li QQ, W YQ, Murphy, Robert W, Blair Christopher, Wu SF, Yue BS, Zhang, YP. 2013. Ecological genetics of Chinese rhesus macaque in response to mountain building: all things are not equal. *PLoS One*, 8(2): e55315. (doi: 10.1371/journal.pone.0055315)

Yang Z, Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple Loci. *Molecular Biology Evolution* 31:3125-3135. (doi: 10.1093/molbev/msu279)

Ye Z, Chen P, Bu W. 2016a. Terrestrial mountain islands and Pleistocene climate fluctuations as motors for speciation: A case study on the genus *Pseudovelia* (Hemiptera: Veliidae). *entific Reports*, 6:33625. (doi: 10.1038/srep33625)

Ye Z, Zhu G, Damgaard J, Chen Xin, Chen PP, Bu WJ. 2016b. Phylogeography of a semi-aquatic bug, *Microvelia horvathi* (Hemiptera: Veliidae): an evaluation of historical, geographical and ecological factors. *entific Reports*, 6(1):2193. (doi: 10.1038/srep21932)





