Lin et al. *Zool. Res.* 2023, 44(1): 3−19
https://doi.org/10.24272/j.issn.2095-8137.2022.108

ZR | Zoological Research

Article Open Access

A wide hybrid zone mediated by precipitation contributed to confused geographical structure of *Scutiger boulengeri*

Xiu-Qin Lin[1,2], Yin-Meng Hou[1,2], Wei-Zhao Yang[1,2], Sheng-Chao Shi[1,2], Pu-Yang Zheng[1,2], Chung-Kun Shih[3,4], Jian-Ping Jiang[1,2,5], Feng Xie[1,2,5,*]

[1] *CAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization and Ecological Restoration Biodiversity Conservation Key Laboratory of Sichuan Province*, *Chengdu Institute of Biology*, *Chinese Academy of Sciences*, *Chengdu*, *Sichuan* 610041, *China*

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

[3] *College of Life Sciences*, *Capital Normal University*, *Beijing* 100048, *China*

[4] *Department of Paleobiology*, *National Museum of Natural History*, *Smithsonian Institution*, *Washington DC* 20013–7012, *USA*

[5] *Mangkang Biodiversity and Ecological Station*, *Xizang Ecological Safety Monitor Network*, *Changdu*, *Xizang* 854500, *China*

ABSTRACT

Confused geographical structure of a population and mitonuclear discordance are shaped by a combination of rapid changes in population demographics and shifts in ecology. In this study, we generated a time-calibrated phylogeny of *Scutiger boulengeri*, an endemic Xizang alpine toad occurring in mountain streams on the Qinghai-Xizang (Tibet) Plateau (QTP). Based on three mitochondrial DNA (mtDNA) genes, eight clades were assigned to three deeply divergent lineages. Analysis of nuclear DNA (nuDNA) genes revealed three distinct clusters without geographic structure, indicating significantly high rates of gene flow. Coalescent theory framework analysis (approximate Bayesian computation model DIYABC and Migrate-N) suggested that divergence of the main intraspecific clusters was the result of hybridization after secondary contact in the Holocene around 0.59 million years ago (Ma). The ratio of mtDNA F_{ST} (fixation index) to nuDNA F_{ST} was 2.3, thus failing to show male-biased dispersal. Geographic cline analysis showed that a wide hybrid zone was initially established in southwestern China, without significant reproductive isolation but with strong introgression in *S. boulengeri*, suggesting high hybrid fitness. Furthermore, mtDNA genes exhibited isolation by distance (IBD) while nuDNA genes exhibited significant isolation by environment (IBE). Results suggested that mitonuclear discordance may have initially been caused by geographic isolation, followed by precipitation-mediated hybridization, producing a wide hybrid zone and geographic structure confusion of nuDNA genes in *S. boulengeri*. This study indicated that complicated

Received: 07 June 2022; Accepted: 19 September 2022; Online: 20 September 2022

Foundation items: This study was supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP, 2019QZKK05010503), Biodiversity Survey and Assessment Project of the Ministry of Ecology and Environment, China (2019HJ2096001006), Construction of Basic Conditions Platform of Sichuan Science and Technology Department (2019JDPT0020), and China Biodiversity Observation Networks (Sino BON)

*Corresponding author, E-mail: xiefeng@cib.ac.cn

historical processes may have led to specific genetic patterns, with a specific climate factor facilitating gene flow in the system.

Keywords: Hybrid zone; Mitonuclear discordance; Precipitation; Genetic structure; Secondary contact

INTRODUCTION

Quantifying the population genetic structure to interpret and predict hybrid zone status can provide potential insights into how demographic history and gene flow change with habitat alteration (López-Uribe et al., 2019). Niche models focus on predicted changes in abiotic factors, while population dynamics also explore internal drivers that influence hybridization or selection in the face of changing climate (Taylor et al., 2015). Although our understanding of the environmental and ecological contexts has increased, how complex population processes, such as genetic admixture and demographic history, contribute to emergent conflicts in mitonuclear evolutionary patterns remains unclear (Parins-Fukuchi et al., 2021).

Mitonuclear discordance in multilocus genomic datasets can be caused by evolutionary processes such as sex-biased dispersal (Prugnolle & de Meeus, 2002), asymmetrical introgression (Yang et al., 2021), incomplete lineage sorting (Wang et al., 2018), and natural selection (Toews & Brelsford, 2012). Despite advances in phylogeny and its emphasis on historical causation, many empiricists remain strongly focused on recurrent gene flow as an explanation for population genetic structure, without considering the demographic history hidden within the genealogy of sampled alleles (Marko & Hart, 2011). Changes in genetic diversity across species-level distribution ranges is highly informative for assessing recent demographic changes and local adaptation potential (Major et al., 2021). Population structure is often used as an indirect measure of gene flow, providing partial insights into species dispersal and colonization ability (Bohonak, 1999; Broquet & Petit, 2009), and is usually measured by a combination of genetic markers (DNA sequences) and theoretical models (Wilson & Rannala, 2003). Various software programs exist for simulating datasets under coalescent theoretical frameworks as well as inferring parameters, such as population size and migration rates, from genetic data as indirect estimators of gene flow (Cornuet et al., 2014; Drummond et al., 2012), However, understanding population history remains challenging, especially population demographic events such as rapid fluctuations in population size and genetic exchange between diverging populations with continuous introgression and local adaptations (Blanquart et al., 2012; Twyford et al., 2020).

Hybrid zones, i.e., geographic regions where genetically distinct groups of individuals interact and produce offspring of mixed ancestry, have long been recognized as tractable windows into evolutionary processes (Barton & Hewitt, 1989; Hewitt, 1988; Larson et al., 2014). Both direct and indirect impacts of climate change can affect spatiotemporal dynamics of hybrid zones (McQuillan & Rice, 2015; Menon et al., 2018). Geographic cline analysis can identify population structures formed by hybridization or selection by plotting changes in allele frequencies along a geographic transect across a hybrid zone (Derryberry et al., 2014). Observed changes in allele frequency at the focal locus are then plotted against expected change based on the mean hybrid index, with significant deviations suggesting that genes are under selection (Szymura & Barton, 1991). Individual movement is thought to reduce differences in mean phenotype among populations, slowing or even preventing local adaptation (Blanquart et al., 2012) but promoting admixture if reproductive isolation has not yet been established. Hybrid zone dynamics may also be driven by strong selection for novel phenotypes, and both phenotype and gene flow can vary due to changes in the environment (Boaratti & Da Silva, 2015). Thus, studying phenotypic changes in hybrid zones can help elucidate the complex mechanisms underpinning niche range shift by internal genetic structure or external drivers of phenotypic changes in the system.

Scutiger boulengeri, an endemic cold-adapted Xizang alpine toad occurring in mountain streams in Xizang (Tibet) (Hofmann et al., 2017), is widely distributed along the eastern and southern slopes of the Qinghai-Xizang Plateau (QTP) at elevations between 2 400 and 5 270 m a.s.l. (Chen et al., 2009; Lin et al., 2021). Previous research of this taxon has revealed that three out of six distinct mitochondrial DNA (mtDNA) genetic clades exhibit significant niche divergence and corresponding body size variation (Lin et al., 2021). This system facilitating exploration of novel niche space may work synergistically with intrinsic and extrinsic factors driving population structure and the emergence of new phenotypic changes under climate change.

Here, we used multilocus mitochondrial DNA (mtDNA) and nuclear DNA (nuDNA) genes combined with demographic models and statistical analyses to clarify the population evolutionary processes underlying alpine toad divergence to characterize their genetic structure and the contributions of demographic processes and phenotypic variation. We also explored the following questions: (I) Do mtDNA and nuDNA genes show discordance? (II) How do population processes and/or heterogeneous environments contribute to the distinct clusters? (III) Do phenotypic characteristics play a role in mediating the population structure of *S. boulengeri*? (IV) Do geographic isolation, hybridization, and environmental selection jointly contribute to such a system? We anticipate that differences in phenotypic traits associated with sex-biased dispersal and population structure may contribute to differences in population connectivity in widely dispersed Xizang alpine toads. This work paves the way for revealing the historical processes that may have led to these patterns of genetic and phenotypic diversity.

MATERIALS AND METHODS

Occurrence, environmental, and phenotypic data

In total, 24 study sites were sampled for *S. boulengeri* during field work conducted from 2019–2021. Localities covered the Himalayas, QTP, Hengduan Mountains, Min Mountains, and adjacent mountains, as per Lin et al. (2021). Detailed

information is shown in Supplementary Table S1. All animal handling and processing were conducted in accordance with the Laws of the People's Republic of China on the Protection of Wildlife and were approved by the Animal Care Committee of the Chengdu Institute of Biology, Chinese Academy of Sciences.

We used 19 BIOCLIM variables and elevation with a resolution of 30 s (1 km) for each environmental layer in WorldClim v2 (Fick & Hijmans, 2017). We examined pairwise correlations among the 20 variables and reduced autocorrelation of input environmental data by removing highly correlated variables with a Pearson correlation threshold of $|r|>0.8$ (Fick & Hijmans, 2017). We selected eight variables with low correlation for subsequent analyses.

We captured 158 adult males (identified from two pairs of keratinized spine patches on the chest) at 19 occurrence sites and obtained morphological measurements following Lin et al. (2021). All morphological variables are shown in Supplementary Table S2. We retained snout-vent length (SVL) for subsequent analyses as it was highly correlated with other morphological variables (see Supplementary Figure S1).

DNA extraction and genetic sequencing

Based on three years of field work, 269 *S. boulengeri* tissue samples were collected. Adult males were identified based on morphological characteristics, while immature individuals and females were identified based on mitochondrial genes. All tissue samples were stored at −20 °C in 96% ethanol.

For mitochondrial cytochrome *b* (cyt *b*), cytochrome c oxidase subunit I (*COI*), and *16S* ribosomal RNA (*16S*) gene fragments, polymerase chain reaction (PCR) amplification was performed with annealing temperatures of 50 °C, 56 °C, and 46 °C, respectively (Supplementary Table S3). Introns can provide sufficient information for population-level questions. Standard PCR amplification was performed at an annealing temperature of 59 °C for cellular myelocytomatosis gene intron 2 (*cmyc-2*, Chen et al., 2009). Exons can provide appropriate levels of variation, ease of alignment, and simple detection of paralogs (Shen et al., 2013). PCR was performed at annealing temperature of 56 °C for the *RAG1* gene. Two-round (nested) PCR was adopted for other exons. The first and second-round PCR procedures were performed in a 25 μL reaction system, with cycling conditions following Shen et al. (2013). We harvested three newly obtained mtDNA genes: i.e., *16S* (534 bp), *COI* (645 bp), and cyt *b* (990 bp) (203 samples in total, including 171 *S. boulengeri* samples from 24 localities, nine *S. glandulatus* samples from three localities, 12 *S. mammatus* samples from three localities, and 11 *S. nyingchiensis* samples from one locality; Figure 1C, GenBank accession Nos. for *16S*: OM505294–OM505496; *COI*: OM505977–OM506179; cyt *b*: OM505096–OM505293 and MW600725–MW600729. NucBank accession Nos. for *16S*: C_AA000899.1–C_AA001101.1; *COI*: C_AA000576.1–C_AA000778.1; *cyt b*: C_AA000373.1–C_AA000575.1) and four nuDNA genes, including one intron (*cmyc-2*, 1 195 bp) and three exons (*DOLK*, 720 bp; *KIAA*, 554 bp; *RAG1*, 909 bp) (*n*=120 from 18 localities, Table 1, GenBank accession Nos. for *cmyc-2*: OM505497–OM505616; *DOLK*: OM505617–OM505736; *KIAA*: OM505737–OM505856; and *RAG1*: OM505857–OM505976. NucBank accession Nos. for *cmyc-2*: C_AA000253.1–C_AA000372.1; *DOLK*: C_AA000013.1–C_AA000132.1; *KIAA*: C_AA000779.1–C_AA000898.1; and *RAG1*: C_AA000133.1–C_AA000252.1).

Phylogenetic reconstruction based on mtDNA

For phylogenetic construction of *S. boulengeri* clades, we used MEGA-X (Kumar et al., 2018) to align all mtDNA fragments, with outgroups *Oreolalax omeimontis* (GenBank IDs for individual KIZ-O.O96002, cyt *b*: EU180928; *16S*: EU180886; *COI*: absent) and *S. nyingchiensis*. Phylogenetic trees were constructed using maximum-likelihood (ML) and Bayesian inference (BI) implemented in PhyloSuite v1.2.1 (Zhang et al., 2020). Best-fit Bayesian Information Criterion (BIC) substitution models (TVM+I+G for *16S*; TVM+I+G for *COI*; GTR+I+G for cyt *b*) were selected using ModelFinder (Kalyaanamoorthy et al., 2017). Divergence times of the phylogenies were estimated using BEAST v1.8.2 (Drummond et al., 2012). Based on previous research, we chose the most recent common ancestor (MRCA) of *Scutiger* and *Oreolalax* (53 million years ago (Ma)) as the calibration point (Hofmann et al., 2017). We performed Markov Chain Monte Carlo (MCMC) simulations for 200 million generations, with the tree sampled every 2 000 generations. For each tree statistic in Tracer v1.7 (Rambaut et al., 2018), effective sample size (ESS>200) indicated satisfactory convergence of the Bayesian chain and adequate model mixing. A maximum clade credibility tree was generated with mean node heights and 10% burn-in using TreeAnnotator v1.8.2 (Drummond et al., 2012). Consensus topologies, branch lengths, and bipartition posterior probabilities were estimated from the remaining trees, with posterior probabilities ≥95% considered as significant. Both the ML and BI trees were visualized and edited in FigTree v1.4.3 (Rambaut & Drummond, 2012).

Population structure and sex-biased dispersal

We converted FASTA files to STRUCTURE files following Banta (2020) and analyzed multilocus nuDNA gene datasets using STRUCTURE v2.3.4 (Pritchard et al., 2000) to determine population structure across the distribution range. As the history of any population is *a priori* unknown and may result from ancient or recent divergence of all populations from a common ancestral population, we ran an admixture model without location prior. In order to quantify the amount of variation of the likelihood for each *K*, we used consecutive *K* values of 1–10, burn in of 5 000, and Markov chain length of 100 000, with three iterations per *K* value. To identify the optimal value for *K*, we calculated mean likelihood L(*K*) and Δ*K* following Evanno et al. (2005) via the pophelper and gridExtra R packages (Auguie & Antonov, 2017; Francis, 2017), as shown in Supplementary Figure S2.

We used ARLEQUIN v3.5.1 (Excoffier et al., 2009) to perform hierarchical analysis of genetic differentiation and variation and analysis of molecular variance (AMOVA) of mtDNA and nuDNA genes between populations using pairwise F_{ST} across loci with 100 000 permutations. Markov chain steps were set to 100 000 and burn-in steps were set to 10 000. The significance of fixation indices was tested using

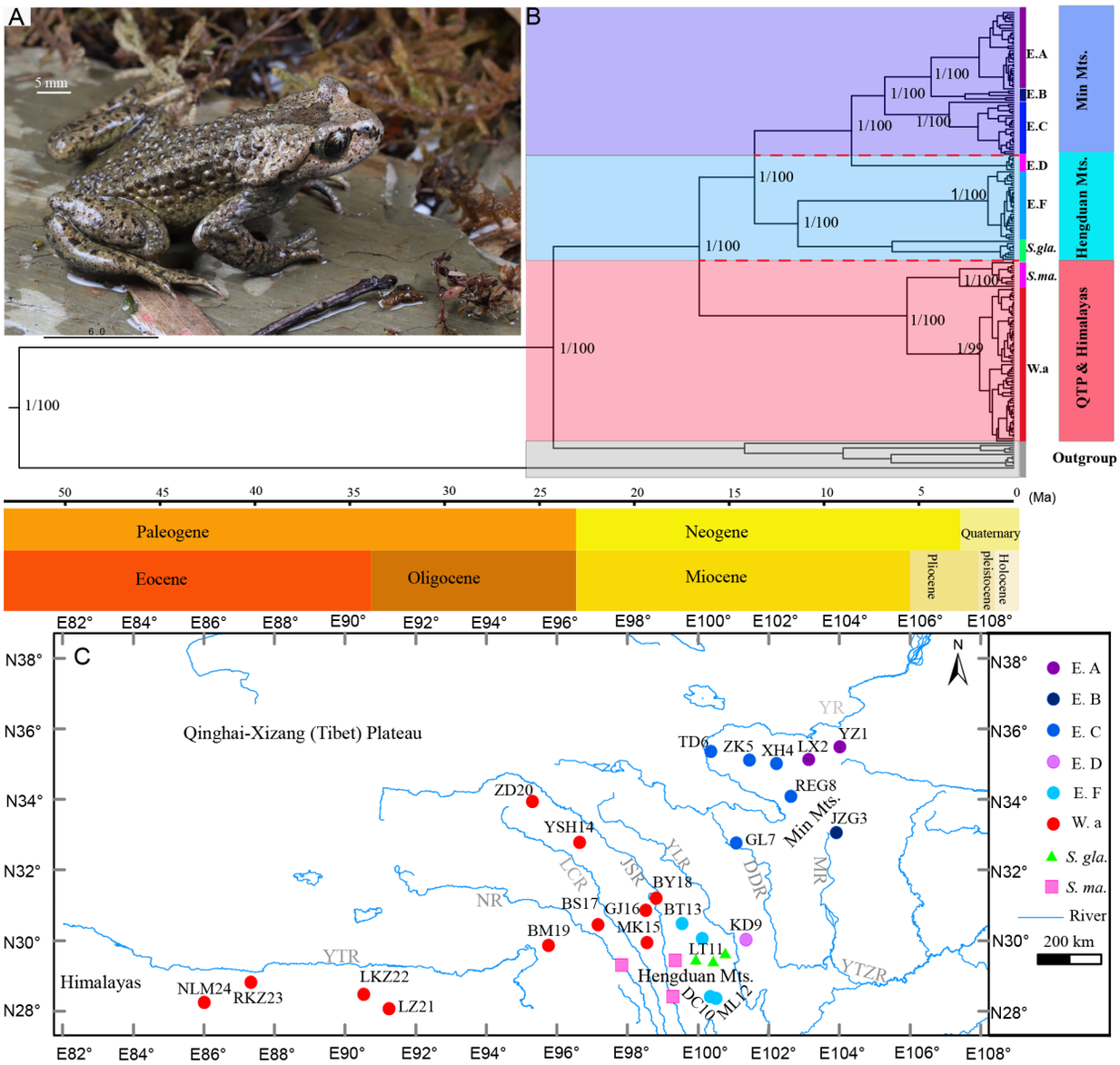

Figure 1 Time-calibrated phylogeny based on mitochondrial DNA and distribution of sampling sites

A: Photograph of focal species *Scutiger boulengeri* (Sample ID: REG20210806, male). B: Time-calibrated phylogeny based on mtDNA genes (*16S*+*COI*+cyt *b*) in *S. boulengeri*, red dashed lines represent main rivers potentially isolating eight clades into three lineages. Ma: Million years ago; Mts.: Mountains; QTP: Qinghai-Xizang (Tibet) Plateau; *S. gla*.: *Scutiger glandulatus*; *S. ma*.: *Scutiger mammatus*. C: Distributions based on occurrence records for *S. boulengeri* clades and related species *S. glandulatus* and *S. mammatus*. Map showing QTP-related regions, including main mountains and rivers. YTR: Yarlung-Tsangpo River; NR: Nu River; LCR: Lancang River; JSR: Jinsha River; YLR: Yalong River; DDR: Dadu River; MR: Min River; YTZR: Yangtze River; YR: Yellow River; Mts.: Mountains; YZ1: Yuzhong (A); LX2: Linxia (B); JZG3: Jiuzhaigou (D); XH4: Xiahe (F); TD5: Tongde (H); ZK6: Zeku (G); GL7: Guoluo; REG8: Ruoergai (E); KD9: Kangding; DC10: Daocheng; LT11: Litang (J); ML12: Muli (K); BT13: Batang (L); YSH14: Yushu (P); MK15: Mangkang (M); GJ16: Gongjue (N); BS17: Basu (O); BY18: Baiyu; BM19: Bomi; ZD20: Zaduo; LZ21: Luozha; LKZ22: Langkazi; RKZ23: Rikeze; NLM24: Nielamu.

10 000 permutations. Due to the different modes of inheritance of mtDNA genes (matrilineal inheritance) and nuDNA genes (biparental inheritance), the divergence ratio of mtDNA to nuDNA genes should be close or equal to 4 (Hudson & Turelli, 2003; Urquhart et al., 2009). Following the method proposed by Crease et al. (1990), we calculated the divergence ratio of mtDNA to nuDNA genes, with a ratio greater than 4 indicating male-biased dispersal.

Discriminant analysis of principal components (DAPC)

To assess the relationships between different clusters, an adequate method should focus on between-group variation, while neglecting within-group variation (Jombart et al., 2010). DAPC relies on data transformation using principal component analysis (PCA) as a previous step to DA, which ensures that variables for DA are perfectly uncorrelated to minimize variation within clusters, while avoiding over-fitting and

Table 1 Nuclear gene sample information of *Scutiger boulengeri*

Site ID	Location	Province	Latitude (°N)	Longitude (°E)	Altitude (m, a.s.l.)	Sample ID
A	Yuzhong	Gansu	35.48745	104.0101	2 800	A01−A13 (13)
B	Linxia	Gansu	35.1359	103.1416	2 770	A14−A26 (13)
C	Zhuoni	Gansu	34.27166	103.262	3 133	A27−A34 (8)
D	Jiuzhaigou	Sichuan	33.25759	104.2495	2 900	B35 (1)
E	Ruoergai	Sichuan	34.0827	102.6256	3 392	C36−C44 (9)
F	Xiahe	Gansu	35.01339	102.2149	3 170	C45−C53 (9)
G	Zeku	Qinghai	35.11549	101.4652	3 429	C54−C65 (12)
H	Tongde	Qinghai	35.35801	100.3658	3 451	C66−C75(10)
I	Yajiang	Sichuan	30.02772	101.3691	4 269	D76−D81 (6)
J	Litang	Sichuan	30.06709	100.1173	3 999	E82−E87 (6)
K	Muli	Sichuan	28.3592	100.5184	4 325	E88−E93 (6)
L	Batang	Sichuan	30.48216	99.5467	2 542	E94−E102 (9)
M	Mangkang	Xizang	29.94484	98.54653	4 187	F103 (1)
N	Gongjue	Xizang	30.87244	98.52251	4 010	F104−F106 (3)
O	Basu	Xizang	30.45389	97.1726	4 256	F107−F109 (3)
P	Yushu	Qinghai	32.51382	96.57492	3 927	F110 (1)
Q	Zhiduo	Qinghai	33.89308	95.33563	4 447	F111−F117 (7)
R	Bomi	Xizang	29.866244	95.772111	2 804	F118−F120 (3)

instability when using too many PCs (Jombart & Collins, 2015). Based on BIC, we assessed the best-supported model corresponding to the optimal number of clusters using the find.clusters function in the adegenet package in R v.4.0.0. Ideally, the optimal clustering solution should correspond to the lowest BIC (Jombart & Collins, 2015). We used DAPC to investigate the genetic structure of the 18 sampled populations (120 individuals with 3 286 loci). We retained 100 PCs for PCA during preliminary variable transformation, which accounted for more than 90% of total genetic variability. Examination of BIC values from 1 to 100 clusters clearly showed that a subdivision into four clusters should be considered, compared to the three clusters suggested by STRUCTURE. We then investigated the genetic structure of the four newly inferred groups using DAPC and found two clusters closely related in PCA position. Thus, we retained the optimal *K* value suggested by STRUCTURE and performed the following analyses.

Phylogenetic-network analysis

Given the high level of admixture among clusters, we visualized evolutionary relationships by considering the possibility of recombination within or between lineages using the phylogenetic Neighbor-Net algorithm implemented in SplitsTree4 v4.12 (Bryant & Moulton, 2004). SplitsTree4 v4.12 analysis was based on aligned FASTA format, as recommended by Huson & Bryant (2008). Recombination, a powerful evolutionary force merging historically distinct genotypes, was tested using the pairwise homoplasy index (Phi) (Bruen et al., 2006; Huson & Bryant, 2006) with the "conduct Phi test for recombination" function in SplitsTree4 v4.12 (Bryant & Moulton, 2004). In addition, we identified recombination sites by counting single crossover (Huson & Kloepper, 2005).

Tests of migration and population size

We used coalescence-based Migrate-N v4.4.4 (Beerli et al., 2019) to explicitly test population structure and gene flow in *S. boulengeri*. We compared different population models, then ranked them according to their marginal likelihood and calculated Bayes factors (BF=0 for the highest-ranking models). Each Migrate-N analysis was run for 50 million steps and recorded every 10 steps, with a burn-in of 5 million and default heating scheme. Suitable upper bounds for population size (Θ) and migration rate (M) priors were determined from an initial test run of 10 million steps per analysis.

Population history by approximate Bayesian computation (ABC) frame modeling

Despite potential gene flow, we used K=3 in downstream analyses. ABC was performed for nuDNA sequence data, as implemented in DIYABC v2.1.0, to explore population history using a coalescence-based method (Cornuet et al., 2014). Based on linear discriminant analysis of summary statistics prior to logistic regression analysis, DIYABC allows for complex population histories to be considered, including any combination of population divergence events, admixture events, and past population size changes (Cornuet et al., 2014). DIYABC generates a simulated dataset, which is then used to select those most similar to the observed and selected datasets, and finally estimate the posterior distribution of parameters (Cornuet et al., 2008, 2014).

We constructed three scenarios to simulate population structure dynamics. All STRUCTURE clusters were treated as distinct populations and 1 000 000 simulated datasets were run for each scenario. We compared three hypothetical models of speciation by analyzing the nuDNA gene datasets using DIYABC v2.1.0 (Cornuet et al., 2014). In all three models, t_d was the divergence time between ancestral populations of *S. boulengeri*, and N_s was the effective population size of the common ancestor of the focal species; t_a was the secondary contact time between two initially divergent populations of *S. boulengeri*, r_a was the section from lineage (cluster) 2, and $1–r_a$ was the section from lineage

(cluster) 3. Model 1 (common origin followed by isolation): three parental clusters (1, 2, and 3) with constant effective population sizes N_1, N_2 and N_3 and divergence time t_d from ancestral population of size N_s. Model 2 (incomplete lineage sorting; ILS): cluster 1 is derived from cluster 2, which is derived from cluster 3. Model 3 (secondary contact): two parental clusters (2 and 3) with constant effective population sizes N_2 and N_3 and divergence time t_d from ancestral population of size N_s. At time t_a, an admixture event between two clusters occurred, giving birth to an admixed cluster with effective size N_1 and admixture rate r_a relative to cluster 2. All prior distributions for model parameters used in the divergence model comparisons are shown in Table 2.

Geographic cline analyses

Sigmoid clines were fit to the hybrid index (based on STRUCTURE Q scores) along geographic transects spanning the phylogeographic transitions using the R package HZAR (Derryberry et al., 2014). One-dimensional sampling of the hybrid zone assumes minimal variation perpendicular to the cline (Derryberry et al., 2014). Here, geographic distance was determined using the geosphere R package (Hijmans et al., 2021). We tested different cline models available in HZAR that can fit molecular data ranging from two (width w and center c) to eight parameters, including constrained or unconstrained exponential tails (Brumfield et al., 2001), and selected those with the lowest sample size-corrected Akaike information criterion (AICc, Dufresnes et al., 2020). The estimated parameters were then used to determine concordance and coincidence of clines from different loci and infer strength of selection against hybrids (Derryberry et al., 2014).

Landscape genetic analysis

How geographic and environmental variables affect neutral and adaptive genetic variations in natural populations is a key question in evolutionary biology (Medina et al., 2021). To identify potential drivers of population genetic structure, we used multiple matrix regression with randomization (MMRR), partial Mantel tests, and Mantel tests to estimate correlations between multiple matrices related to genetic, environmental, and geographic distances. We obtained the top five PC loadings and climate variable contributions based on the PCA function in the FactoMineR package in R v4.0.0. Environmental distances (Euclidean) between sites were calculated from variables with a Pearson correlation threshold of $|r|<0.8$. Pairwise geographic distances between sampled populations were calculated using the distm function (Hijmans et al., 2021) according to the "Haversine method" in the geosphere package of R v4.0.0. Environmental distances were determined using the dist function in the ade4 package. When gene flow is potentially influenced by multiple explanatory variables, such as geographic and environmental distances, MMRR provides a valuable tool to describe how genetic distances respond to these variables (Wang, 2013). To identify the variables contributing to the optimal model, we performed stepwise regression elimination until only variables with significant *P*-values remained.

To exclude the spatial effects proposed by Bonnet & Van de Peer (2002), we estimated the correlations among geographic factors, environmental factors, and F_{ST} using partial Mantel tests (Legendre & Fortin, 2010). Partial Mantel tests involve three matrices, with the aim to test the correlation between matrices A and B while controlling the effect of matrix C to remove spurious correlations (Bonnet & Van de Peer, 2002; Legendre & Fortin, 2010). Here, partial Mantel tests were performed between genetic distances and one factor under the influence of the other (as a covariate) using zt v1.1 (Bonnet & Van de Peer, 2002) with the command "zt.exe -p file1.txt file2.txt file3.txt 10000" implemented in the dos box. Mantel tests were performed between genetic distances and other factors using the command "zt.exe -s file1.txt file2.txt 10000" implemented in the dos box.

Table 2 Prior distributions for model parameters used in divergence model comparisons

Parameter	Scenario		
	Prior distribution	Minimum	Maximum
Effective population size			
N_s	Uniform	10 000	1 000 000
N_1	Uniform	10 000	1 000 000
N_2	Uniform	10 000	1 000 000
N_3	Uniform	10 000	1 000 000
Time of events			
t_a	Uniform	1 000	1 000 000
t_d	Uniform	10 000	1 000 000
Admixture rate			
r_a	Uniform	0.001	0.999
Mean mutation rate uniform	Uniform	1.0×10^{-9}	1.0×10^{-7}
Individual locus mutation rate	Gamma	1.0×10^{-8}	1.0×10^{-6}
Mean coefficient K_C/T	Uniform	0.05	0.9
Individual locus coefficient K_C/T	Gamma	0.05	0.9

N_1, N_2, and N_3 represent constant effective population sizes of three parental lineages (clusters 1, 2, and 3), respectively; N_s is effective population size of common ancestor of focal species; t_a is secondary contact time between two initial divergent populations of *S. boulengeri*, t_d is divergence time between ancestral populations of *S. boulengeri*; r_a is section from lineage (cluster) 2, and 1-r_a is section from lineage (cluster) 3.

RESULTS

Phylogenetic reconstruction based on mtDNA

After manual sequence correction, alignment, and end trimming, concatenated gene sequences (2 169 bp) were obtained with three fragments, i.e., *16S* (534 bp), *COI* (645 bp), and cyt *b* (990 bp), from 203 samples (including *S. glandulatus*, *S. mammatus*, and *S. nyingchiensis*). No gaps were found, and all sequences were correctly translated into amino acid sequences based on vertebrate genetic codes. The best substitution models (TVM+I+G for *16S*, TVM+I+G for *COI*, and GTR+I+G for cyt *b*) were identified using PartitionFinder v2.1.1. Both BI and ML phylogenetic trees clearly divided *S. boulengeri* into three deeply split lineages (i.e., Min Mts. lineage, Hengduan Mts. lineage, and Qinghai-Xizang (Tibet) Plateau (QTP) & Himalayas lineage) with at least eight clades (clades E.A, E.B, E.C, E.D, E.F, and W.a, with *S. glandulatus* and *S. mammatus* nested, Figure 1B, C). Analysis indicated that *S. boulengeri* diverged from *S. nyingchiensis* in the early Miocene ca. 23.57 Ma (95% highest probability distribution (HPD): 22.58–25.99 Ma), and the earliest QTP lineage of *S. boulengeri* and *S. mammatus* diverged from ancestral *S. boulengeri* in the Miocene ca. 16.59 Ma (95% HPD: 14.58–18.46 Ma).

Population structure analysis

We assembled one intron and three exons to align the 3 286 bp nuDNA fragments. Description of the four steps allowing detection of the true number of clusters is shown in Supplementary Figure S2. We found that the likelihood L(*K*) value increased until the real *K* (*K*=3) was reached, after which it plateaued (Supplementary Figure S2A), and similarly, the distribution of ΔK showed a maximum at the real *K* value (Supplementary Figure S2D). ΔK reached the highest peak at *K*=3 (Figure 2C, D), indicating that the ancient populations of all individuals were divided into three distinct clusters. These results roughly corresponded to the three main mtDNA lineages. The STRUCTURE results showed relatively high admixture among genetic clusters, with individuals from different regions assigned to the same cluster (Figure 2D), which may be due to recent migration or hybridization. Average F_{ST} in mtDNA genes was 0.748 (varying from −0.162 to 0.982; Supplementary Table S4). Average F_{ST} in nuDNA genes was 0.480 (varying from −0.062 to 1; Figure 2B). Of note, populations D, M, and P each contained only one individual in which the three exon fragments were successfully amplified, thus pairwise F_{ST} values for these populations may have been overestimated. As such, these values were removed when calculating the divergence ratio of mtDNA to nuDNA genes. Here, the divergence ratio was 2.3, much lower than 4, suggesting no obvious sex-biased dispersal.

Discriminant analysis of principal components

The optimal *K* value of STRUCTURE analysis did not match the *K* value corresponding to the lowest BIC value in DAPC, and the optimal *K* value of STRUCTURE analysis was lower than that of DAPC. The *K* value corresponding to the BIC value was slightly smaller, as reported in Everson et al. (2021) (Figure 2C, E). Using the STRUCTURE results as a benchmark for DAPC performance, the color maps showed well-defined patterns that were highly similar to the STRUCTURE results obtained under the four-cluster population genetics model. The eigenvalues indicated that the main genetic structure was captured by the first two PCs in DAPC analysis. The first PCA quadrant (red, Figure 2F) contained cluster 3 from QTP & Himalayas, as suggested by STRUCTURE analysis. The second quadrant (blue, Figure 2F) contained cluster 1 from the Min Mts., and the third quadrant (light blue, Figure 2F) contained cluster 2 from the Hengduan Mts.

Network analysis

Network analysis also recovered groups that generally corresponded to STRUCTURE and DAPC results, but with evidence of many reticulations (Figure 3). Results indicated that part of cluster 1 (i.e., crossover position for recombination at geographic sites A, C, E, D, and I) was the result of recombination between clusters 2 and 3. Furthermore, Phi test analysis identified 355 informative sites in the 3 286 bp nuDNA dataset showing significant recombination, with a mean value for recombination of 0.293 (*P*=0). Bootstrap analysis with 1 000 replicates showed high statistical support. The starburst-like pattern indicated conflicting topologies within the dataset, likely due to biological population dynamics, including recent diversification with frequent gene flow mediated by hybridization and recombination.

Migration and population size

We used Bayesian modeling to test for genetic isolation and gene flow among the three clusters chosen by STRUCTURE analysis. In our study, the top-ranking model was secondary contact with two isolated clusters (Figure 4B). Simulated models of restricted migration and no migration (Figure 4A, C) were rejected with decisive Bayes factors (Table 3). Migrate-N demographic analysis indicated that the origin in *S. boulengeri* formed the N_2 and N_3 clusters and secondary contact formed the N_1 cluster (Figure 4B). The population size of N_1 was 0.48 times that of N_2 and 1.08 times that of N_3. The credibility interval for the size of N_1 was 97.5% HPD: 1.18e-2–2.15e-2, of N_2 was 97.5% HPD: 2.1e-2–4.41e-2, and of N_3 was 97.5% HPD: 9.93e-3–1.99e-2. In addition, under the secondary contact model, migration rates were symmetrical among clusters (Table 4). An important caveat to migration and population size analyses is that divergence time was not included as a model parameter, thus preventing distinction between historical and contemporary gene flow. We performed DIYABC analysis to resolve this issue.

Gene flow and geographic cline analysis

Based on the detection of three clusters (N_1, N_2, and N_3) by STRUCTURE analysis, ABC demographic analysis found that Model 3 (divergent clusters (N_2 and N_3) with secondary contact) outperformed all other migration models (Figures 5C, 6A). N_1 originated from secondary contact between the N_2 and N_3 clusters. The population size of N_1 was about 0.8 times that of N_2 and 1.36 and 6.4 times that of N_3 and N_s, respectively. Population size relationships were similar to the results estimated by Migrate-N at the mutation scale (N_2>N_1>N_3>N_s). First divergence of the focal species that formed clusters N_2

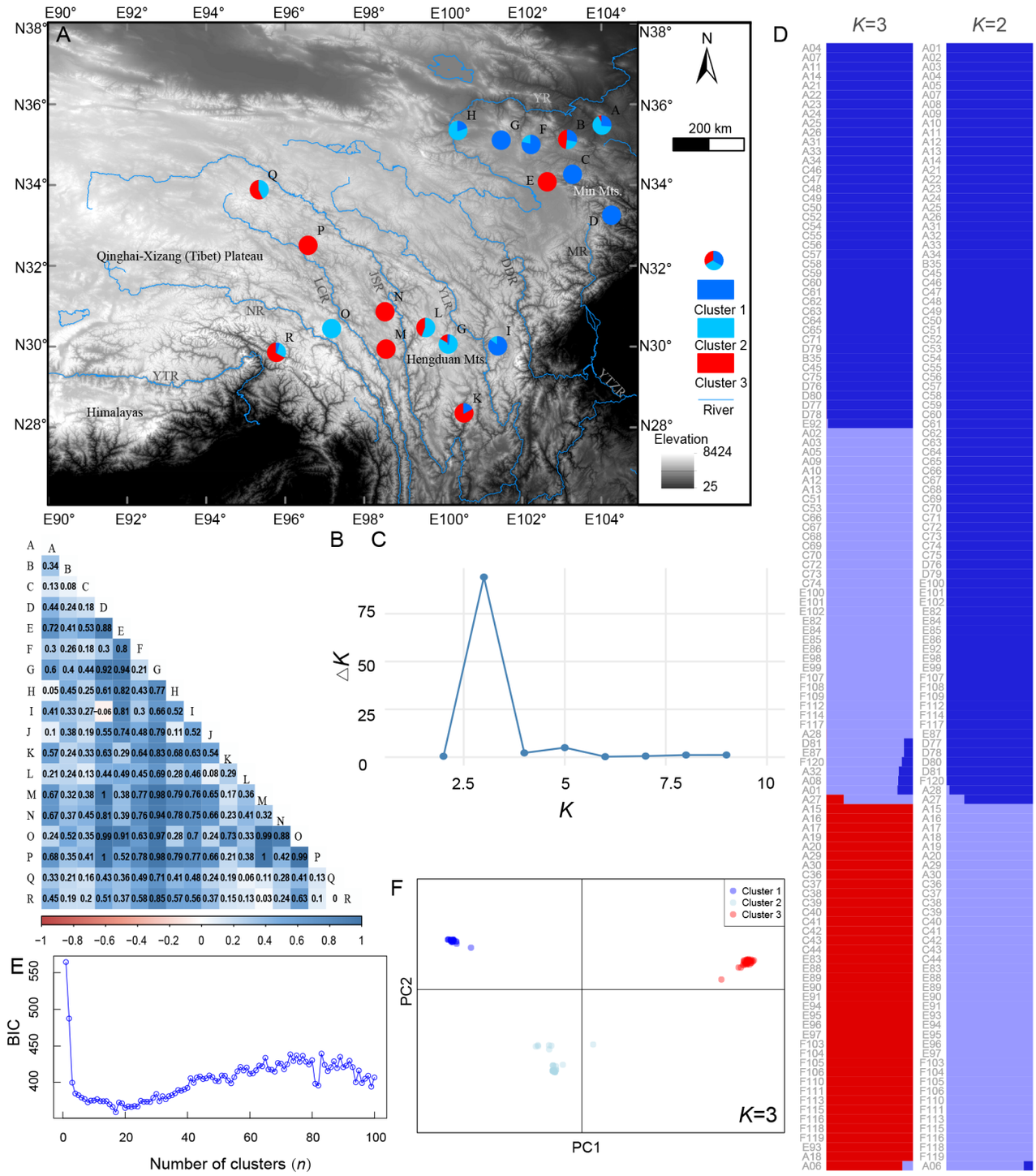

Figure 2 Geographic distribution regions and genetic structure analyses

A: Geographic distribution of 18 sampling localities (A to R) for nuDNA genes of Xizang alpine toads from China. Circles represent proportion of assignment to three genetic clusters identified in C. B: Heatmap of nuDNA F_{ST} distance matrix in pairwise location sites, where darker blue indicates more dissimilar and magenta indicates more similar. C: Modal value of distribution is the true K value or the uppermost level of the structure, with three clusters showing here. D: Individual ancestry coefficient for 120 individuals obtained with STRUCTURE v2.3.4 and based on 3 286 bp nuDNA genes (K=2 and K=3; each vertical line depicts one individual); E: Genetic clusters across study sites inferred using discriminant analysis of principal components (DAPC) with the adegenet package in R. BIC is provided for different numbers of clusters (1 to 100), with the best BIC found for cluster 4; F: Scatterplots of clusters in DAPC.

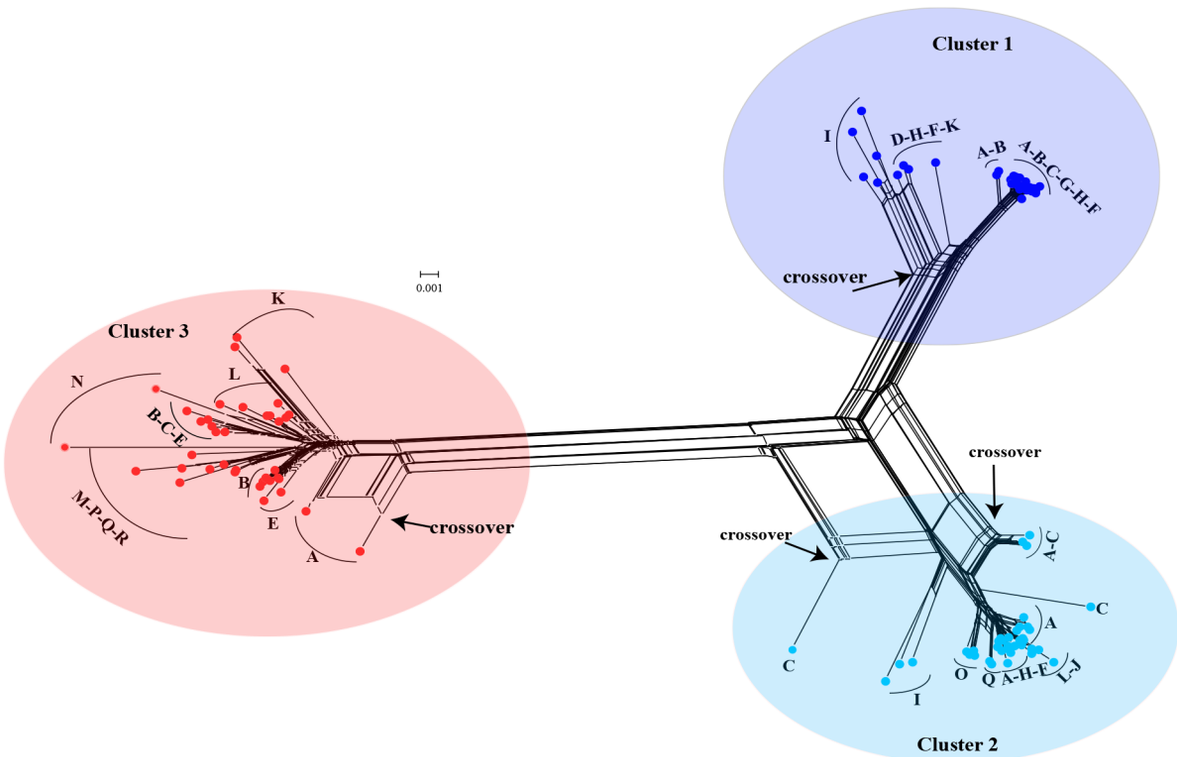

Figure 3 Network analysis implemented in SplitsTree4 v4.12

Networks were constructed using the Neighbor-Net algorithm. Capital letters indicate different geographic locations and colors represent corresponding clusters.

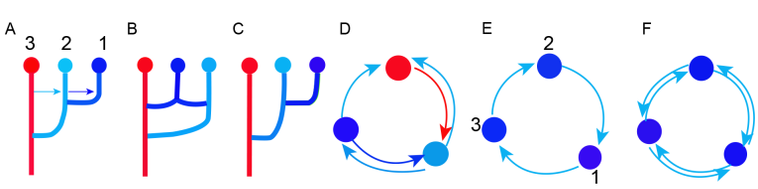

Figure 4 Graphical representation of models evaluated in Migrate-N and models fit using Bayes Factors (Table 2)

A: Geographic isolation model with unidirectional gene flow, where cluster 1 split from cluster 2, and cluster 2 split from ancestral cluster 3 (recurrent immigration after divergence from ancestor to descendant). B: Hybridization after secondary contact model, where some individuals of divergent clusters fused after secondary contact, resulting in a new cluster. C: Geographic isolation model, where cluster 1 split from cluster 2, and cluster 2 split from ancestral cluster 3 without gene flow. D: Three-cluster model, where clusters 1, 2, and 3 exchanged migrants at individual rates with partial bidirectionality. E: Three-cluster model, where clusters 1, 2, and 3 exchanged migrants at individual rates with unidirectionality. F: Three-cluster model, where clusters 1, 2, and 3 exchanged migrants at individual rates with bidirectionality.

Table 3 Natural logarithms of Bézier-corrected marginal likelihoods (Bézier lnL), Bayes factors (BF), and model ranks for each model evaluated in Migrate-N

Model	Log (mL)	BF	Model probability	Rank
A	−14 192.75	−971.45	0	3
B	**−13 221.30**	**0**	**1**	**1**
C	−13 262.48	−41.18	0	2
D	−15 643.54	−2 422.24	0	6
E	−15 557.59	−2 336.29	0	4
F	−15 762.57	−2 541.27	0	5

Top-ranked model is in bold.

and N_3 from the ancestral lineage occurred approximately 497 000 generations ago (97.5% HPD: 152 000–951 000, Table 5). In addition, some individuals from the divergent N_2 and N_3 clusters fused and contributed to northward expansion after secondary contact between the Hengduan Mts. and QTP lineages approximately 118 000 generations ago (97.5% HPD:

Table 4 Parameters estimated from Migrate-N analyses for population size (Θ) and mutation-scaled migration rate

Parameter	2.50%	25.00%	Mode	75.00%	97.50%	Median	Mean
$\Theta 1$	0.01180	0.01207	0.01617	0.02087	0.02153	0.01650	0.01678
$\Theta 2$	0.02067	0.02587	0.02877	0.03380	0.04413	0.03090	0.03119
$\Theta 3$	0.00993	0.01240	0.01417	0.01600	0.01993	0.01470	0.01527
$D_{3\text{-}1}$	0.00000	0.00000	0.00003	0.00113	0.00280	0.00117	0.00913
$S_{3\text{-}1}$	0.00000	0.00000	0.00010	0.00167	0.00300	0.00170	0.00908
$D_{2\text{-}1}$	0.00000	0.00000	0.00003	0.00147	0.00300	0.00150	0.01164
$S_{2\text{-}1}$	0.00000	0.00000	0.00070	0.00307	0.00333	0.01863	0.02064
$D_{3\text{-}2}$	0.00000	0.00000	0.00003	0.00047	0.00147	0.00050	0.00086
$S_{3\text{-}2}$	0.00000	0.00013	0.00063	0.00107	0.00167	0.00090	0.00182

D is divergence from ancestral column population with ongoing immigration; *S* indicates symmetric migration parameter.

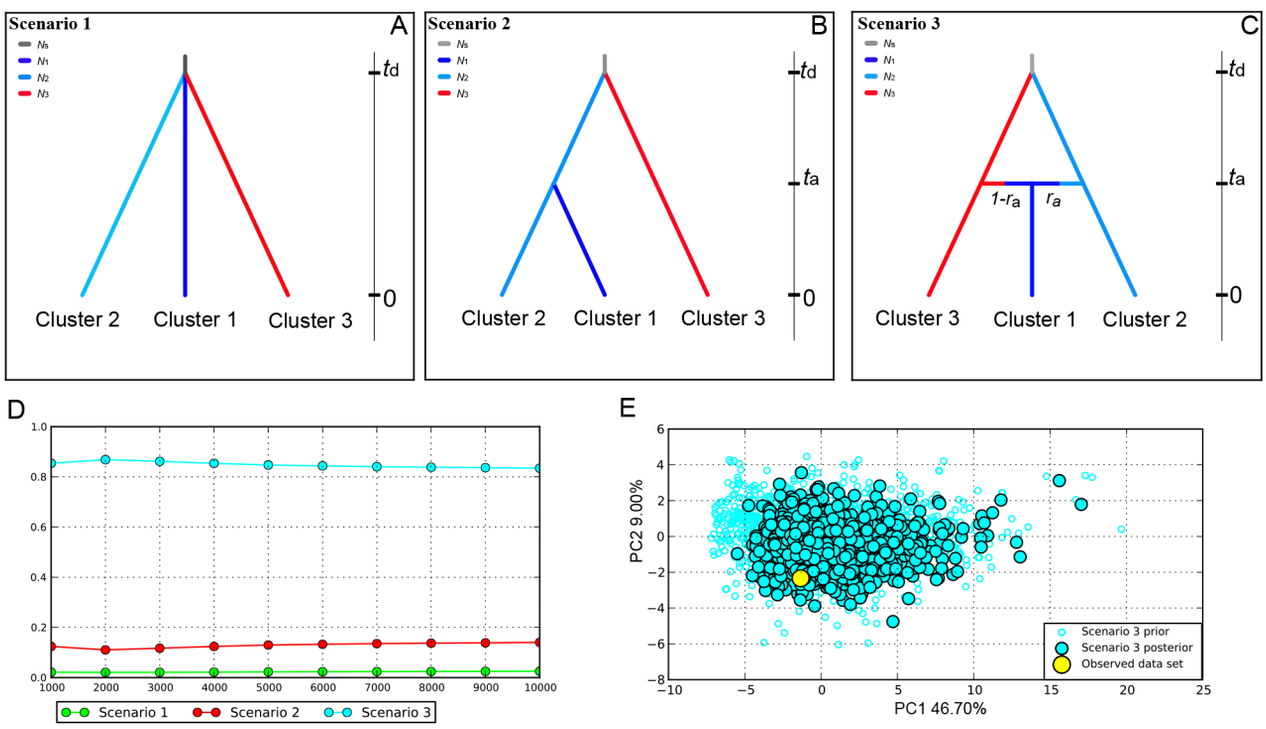

Figure 5 Divergent models based on existing assumptions in *S. boulengeri* and corresponding posterior probabilities for best-selected model in DIYABC analysis

A–C: Illustrations of three alternative models (scenarios). D: Posterior probability of each model was estimated directly and with logistic regression using 0.1% and 1%, respectively, of the simulated data closest to the observed dataset in the model selection procedure. E: Model checking and PCA of observed data, comparing prior and posterior distributions of parameters in model 3.

Table 5 Posterior median, mean, mode, and range of 97.5% highest probability distribution (HPD) for seven demographic parameters in lineage (cluster) merging model for *S. boulengeri* (time is in generations)

	2.5% HPD	Median	Mean	Mode	97.5% HPD
N_1	2.16×10^5	5.06×10^5	5.08×10^5	5.57×10^5	8.46×10^5
N_2	2.87×10^5	6.28×10^5	6.22×10^5	6.24×10^5	9.37×10^5
N_3	1.15×10^5	3.72×10^5	4.02×10^5	2.87×10^5	8.83×10^5
N_s	1.04×10^4	7.88×10^4	7.08×10^4	9.95×10^4	9.93×10^4
t_a	2.96×10^4	1.18×10^5	1.40×10^5	9.44×10^4	4.00×10^5
t_d	1.52×10^5	4.97×10^5	5.14×10^5	4.39×10^5	9.51×10^5
r_a	6.42×10^{-1}	9.81×10^{-1}	9.46×10^{-1}	9.99×10^{-1}	9.99×10^{-1}

t_a is divergence time of N_2 and N_3, r_a is admixture rate of clusters N_2 and N_3. N_1, N_2, N_3, and N_s are effective population sizes of clusters 1–3 and common ancestor of *S. boulengeri*, respectively.

29 600–400 000, Table 5). We calculated divergence time based on the assumption that whole-phase development of an individual requires 5 years per generation (Song et al., 1990). Results showed that divergence of the N_2 and N_3 clusters from the ancestral lineage occurred around 2.49 Ma (97.5% HPD: 0.76–4.75 Ma, Figure 6A), and N_2 and N_3 fused and contributed to northward expansion after secondary contact between the Hengduan Mts. and QTP & Himalayan regions approximately 0.59 Ma (97.5% HPD: 0.14–2 Ma, Figure 6A). The mutation rate (μ) was 5.12e-08 sites/generation (97.5% HPD: 2.30e-08–9.25e-08, Supplementary Figure S3H). The posterior distributions of parameters were located perfectly on the prior interval (Supplementary Figure S3), suggesting that the results were robust. ABC framework demographic analysis showed that the divergence estimates using nuDNA genes joined exons and intron more recently compared to mtDNA genes, indicating that hybridization after secondary contact contributed more to mitonuclear discordance.

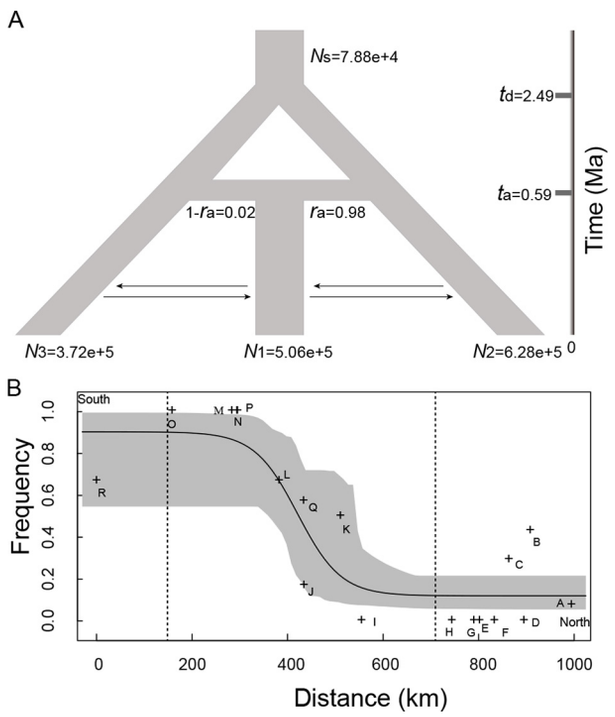

Figure 6 Ideal population historical demographic model and geographic cline analysis

A: Schematic representation of historical population demography and gene flow of *S. boulengeri*. B: Geographic cline analyses of *S. boulengeri* hybrid zone transitions, with 95% credible cline region for selected models based on 3 286 bp nuDNA gene dataset.

Hybrid cline analysis selected Model 4 as the best molecular cline model with the lowest AICc. The hybrid cline center among sites L, Q, K, and J was located 435.46 km around the Hengduan Mts., and the hybrid zone had a cline width of 277.63 km (Figure 6B). Our snapshot showed a wide hybrid cline zone, with its width likely caused by weak selection for hybrids, indicating easy dispersal across the distribution zone and recent gene exchange among populations in the hybrid or even tension zones. A wide and gentle hybrid zone slope with minimal spatial mixing of parental taxa suggests high hybrid fitness.

Landscape genetic analysis

Based on PCA, we obtained the top five PCs, which explained 97.88% of the variance in environmental attributes. The PC1 axis was closely associated with BIO12 (Annual precipitation), BIO7 (Temperature annual range), and BIO14 (Precipitation of driest month), while the PC2 axis was closely associated with BIO3 (Isothermality), BIO8 (Mean temperature of wettest quarter), and BIO2 (Mean diurnal range) (Table 6).

The MMRR analyses quantified the contributions of geographic, genetic, and environmental distances. Results showed that mtDNA F_{ST} differentiation was positively related to geographic distance and negatively related to PC1 (MMRR R^2=0.322, P<0.001), while nuDNA F_{ST} was negatively associated with precipitation differences (considering synergy of multiple precipitation variables), approaching significance (MMRR R^2=0.109, P=0.0439).

To exclude spatial effects, we used the partial Mantel test to explore the correlation between F_{ST} and geographic distance, while controlling for the effect of a third environmental distance matrix. Results showed that geographic distance was highly significantly correlated with mtDNA F_{ST} (partial Mantel r=0.565, P=0.0001) when controlling for the effects of environmental variables along the PC1 axis. These results indicated significant correlation between genetic and

Table 6 Loadings of top five principal components for climate variables with lower correlations

		PC1	PC2	PC3	PC4	PC5
BIO1	Annual mean temperature	0.87	15.287	**36.239**	**8.559**	3.576
BIO2	Mean diurnal range	14.977	**16.983**	2.376	6.952	**20.874**
BIO3	Isothermality	0.005	**27.358**	**21.442**	0.22	6.238
BIO7	Temperature annual range	**21.75**	1.51	10.533	**12.263**	5.806
BIO8	Mean temperature of wettest quarter	8.673	**23.236**	10.326	0.001	0.389
BIO12	Annual precipitation	**24.202**	1.093	1.386	11.815	**19.853**
BIO14	Precipitation of driest month	**15.032**	0.117	**16.409**	**45.82**	9.23
BIO15	Precipitation seasonality	14.492	14.416	1.289	14.371	**34.033**
	Eigenvalue	3.27	2.343	1.561	0.459	0.198
	Percentage of variance	40.878	29.281	19.518	5.732	2.471
	Cumulative variance	40.878	70.159	89.677	95.409	97.88

Loadings of top three are given in bold for each component. PC1 mainly corresponds to precipitation, PC2 mainly corresponds to temperature.

geographic distances, independent from environmental distance. In the nuDNA dataset, when controlling for the effects of precipitation, the PC3-related environmental variables were still highly significantly correlated with F_{ST} (partial Mantel r=−0.229, P=0.023), suggesting that PC3 and precipitation have independent and significant effects on gene flow.

Mantel test analysis of mtDNA revealed a significant positive correlation between genetic and geographic distances (Mantel r=0.528, P=0.0001, Figure 7A), but a non-significant association between genetic and environmental distances (Mantel r=0.0002, P=0.515). In addition, Mantel test analysis showed that nuDNA genetic differentiation was significantly negatively correlated with PC3 (Mantel r=−0.25, P=0.0015, Figure 7B) and Precipitation (centralization of BIO12, BIO14, and BIO15, Mantel r=−0.20, P=0.0012, Figure 7C). SVL distance was significantly negatively correlated with PC1 (Mantel r=−0.18, P=0.035, Figure 7D), especially BIO14 (Precipitation of driest month) (Mantel r=−0.157, P=0.031), suggesting that Precipitation of driest month is a key factor contributing to SVL variation.

The mtDNA and nuDNA gene results identified different models, i.e., mtDNA was impacted by isolation by distance (IBD), while nuDNA genetic differentiation was influenced by isolation by environment (IBE). However, SVL variation was mainly affected by Precipitation of driest month.

DISCUSSION

Significant mitonuclear discordance was found in *S. boulengeri*, and nuDNA gene analysis revealed three distinct clusters with no clear geographic structure. The coalescent theory framework suggested that divergence of the main intraspecific clusters was the result of hybridization after secondary contact in the Holocene (0.59 Ma). Body size was not a significant predictor of gene flow, but IBE was the ideal model for population structure. A wide hybrid zone mediated by precipitation contributed to the confusion in geographic structure.

nuDNA-mtDNA phylogenetic discordance

Previous studies of alpine toad population genetics focused on single mtDNA sequences (cyt *b*) and partially sampled regions, inferring a clear separation between lineages with greatly reduced gene flow (Li et al., 2009). Furthermore, mtDNA clades showed clear geographic patterns, associated with expansion out of the QTP margins into the Min Mts. (Lin et al., 2021). In the current study, based on 2 189 bp multilocus mtDNA fragments, we identified eight completely sorted clades within three lineages. Completely sorted mtDNA lineages provide the advantage of using mtDNA gene trees as estimates of species trees for closely related taxa (Hudson & Turelli, 2003). As mitochondrial phylogeny only provided a limited genealogical view of alpine toad population history, we compiled four nuDNA markers to explore the phylogenetic network relationships and found structural inconsistencies between mtDNA and nuDNA genes. These results indicated that the nuDNA genetic structure was only partially responsive to the three mtDNA lineages, and most genetic relationships were confused (Figure 3). Furthermore, the divergence time of the initial split of clusters occurred 2.49 Ma (0.76–4.75 Ma), about 6.6 times more recent than the age estimate based on the mtDNA gene tree. In addition, mtDNA genes supported the IBD model while nuDNA genes supported the IBE model.

These results are reasonable because mitochondrial genomes are haploid and maternally inherited, effective population size (*Ne*) of mitochondrial loci is generally one-quarter that of nuclear loci, stochastic lineage sorting

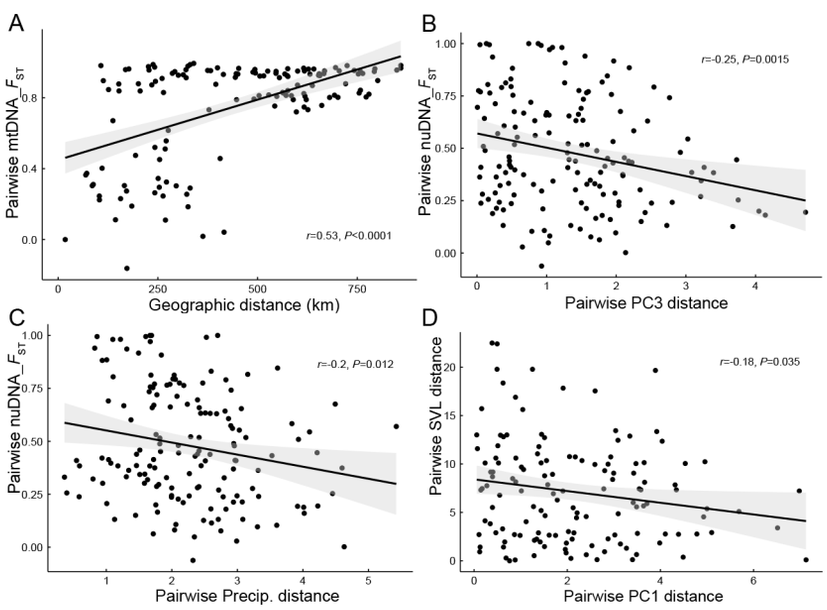

Figure 7 Scatterplots based on Mantel tests

A: Isolation by distance plot using pairwise mtDNA genetic differentiation F_{ST} and geographic distance (great circle distance by Haversine method). B, C: Isolation by environment plots using pairwise nuDNA genetic differentiation F_{ST} and environmental distances (PC3 and Precip.). D: Pairwise SVL distance variation along pairwise PC1 distance. Precip.: precipitation.

progresses more rapidly for mitochondrial alleles (Funk & Omland, 2003), and smaller *N*e of mtDNA compared to nuclear loci may facilitate introgressed haplotype lineages becoming fixed. Després (2019) hypothesized that mtDNA and nuDNA genes respond differently to demographic fluctuations, i.e., when a large population is fragmented into small isolates, such as during glacial periods, mtDNA genes will diverge while nuDNA genes will retain variability. Moreover, lineages are often ephemeral and do not always progress into fully reproductively isolated taxa (Pereira & Singhal, 2022).

Population structure and dynamics

We compiled four unlinked nuDNA genes to test complex demographic scenarios, inferring population differentiation time and gene flow direction. As shown by the nuDNA gene results, clusters did not necessarily respond to "real" geographic populations (Figure 2D). Although an admixture-like structure was detected based on clustering, it was difficult to distinguish real admixtures (resulting from hybridization between divergent clusters) from shared ancestral polymorphisms (inherited from a common ancestor) without using coalescent analysis (Tsuda et al., 2015). Given high coalescent variance, accurate estimation of demographic parameters from gene trees requires multilocus data (Edwards & Bensch, 2009). In this study, the DIYABC and Migrate-N results based on 3 286 bp multilocus nuDNA genes favored the admixture model (Figures 4B, 5C, 6A). Initial divergence occurred around 2.49 Ma in the early Pleistocene, while admixture and recombination occurred following secondary contact between the south (Hengduan Mts.) and west clusters (QTP & Himalayas) around 0.59 Ma during the Holocene. Due to Pleistocene climatic oscillations and regional uplift of orogeny, the temperate Hengduan region is embedded in a network of high mountains surrounding the QTP and is close to cool high-latitude habitats (Xing & Ree, 2017). Regional uplift of orogeny may have contributed more to initial population divergence around 1.2–2.5 Ma (i.e., Xigeda Formation strata, Jiang & Wu, 1998). A consequence of both mountain uplift and increased glaciation is accelerated silicate weathering, which can lead to a reduction in atmospheric CO_2 and global cooling (Head & Gibbard, 2015). Global cooling, which can reduce the strength of the hydrological cycle and increase cold air masses from higher latitudes, may have played a significant role in central Asian aridification (Han et al., 2014). The Naynayxungla Glaciation (0.5–0.78 Ma) had a profound influence on individual dispersal and secondary contact of divided populations of cold-adapted species (e.g., Xizang alpine toads). Regions with distinct physiographic biomes shared the same climatic conditions or were connected by precipitation, which may facilitate gene flow during biogeographic exchange. However, further research with paleoclimate niche models is required.

Wide hybrid zone under a demographic history with continuous gene flow

Geographic cline analysis revealed that the nuDNA genes were characterized by a wide transitional zone, which likely contributed little to reproductive isolation but had a strong influence on introgression in *S. boulengeri*. The wide and gentle slope of the hybrid zone (Figure 6B) suggested high hybrid fitness caused by minimal spatial mixing of parental taxa (Barton & Hewitt, 1985; Buggs, 2007). We speculate that during the Pleistocene climatic oscillations, the Xizang alpine toads moved upward to high-altitude or high-latitude regions during the interglacial periods, leading to isolation by contraction, then moved downward to mid-altitude or mid-latitude regions during the glacial periods, resulting in contact by expansion. The Jinsha and Yalong valleys, two prominent phylogeographic breaks in the Hengduan Mts. (Figures 1C, 2A), showed deep splits between mtDNA clades but no corresponding differences in nuDNA data. According to species distribution modeling, they do not appear to constitute an ecological barrier for *S. boulengeri* (Lin et al., 2021). Geographic cline analysis (Figure 6B, south to north) was similar to the niche simulation results, indicating that Xizang alpine toads rapidly expanded their niche northward (Lin et al., 2021). The northward expansion of a population from southern locations is also found for the green odorous frog (*Odorrana margaretae*) in the Sichuan Basin (Wen & Fu, 2021). In this research, we were unable to infer whether hybrid zone movement occurred, as factors governing such movement are complex, i.e., environmental selection, competition, asymmetrical hybridization, dominance drives, hybrid fitness, human activity, and climate change (Buggs, 2007). Future studies are required to examine clines with more detailed sampling, using whole genomes to connect weakly introgressed loci with traits.

Isolation by environment (IBE) is the ideal model for population structure

Previous research on the genetic structure of *S. boulengeri* populations based on mtDNA genes suggested that geographic barriers, including mountains and rivers, may be key factors affecting gene flow (Li et al., 2009). Lin et al. (2021) proposed that niche divergence caused by northward niche expansion accompanied key morphological innovations of adaptation to novel climates. Our research showed that IBE contributed to the divergent population structure (Figure 7B, C), as reported for the Taliang knobby newt (*Liangshantriton taliangensis*, Shu et al., 2022), suggesting that gene flow rates are higher among similar environments and environmental variables can influence colonization success of individuals and groups via environmental filtering, with higher effective gene movement among similar environments (García-Girón et al., 2019; Sexton et al., 2014).

Body size variation has played an important role in the population genetics of mammals, birds, butterflies, and bees (López-Uribe et al., 2019; Ottaviani et al., 2006; Stevens et al., 2013). In our study, body size of *S. boulengeri* was not a significant predictor of genetic structure, but was negatively correlated to Precipitation of driest month, i.e., larger mean body size was found in drier areas. These results are consistent with the water availability hypothesis, which predicts that larger amphibians can adapt to drier environments because a lower surface/mass ratio reduces water loss (Valenzuela-Sánchez et al., 2015). Our results indicated that Precipitation of driest month (BIO14) occurred in

all three supporting models, with a high independent effect (Mantel r_{BIO14}=−0.157 vs. r_{PC1}=−0.181; P_{BIO14}=0.03 vs. P_{PC1}=0.03), suggesting it may be the most important factor determining variation in body size. This result does not contradict our previously discussed "ecological release" in novel shifted niches (Lin et al., 2021). We propose that synergistic effects may exist between niche expansion, increased body size, and drier environments, with larger individuals expected in drier environments where animals experience long periods without water.

Our study provides insights into how extrinsic environmental factors influence intrinsic population historical demography. Notably, interactions between high dispersal and extrinsic environmental stress (i.e., Precipitation of driest month) may mediate genetic divergence due to environmental adaptation, while, from a genetic spatial aspect, intrinsic factors (i.e., historical demography) may mediate genetic structure. Thus, based on intrinsic genetic mechanisms and extrinsic environmental drivers, precipitation appears to have influenced this complex historical demography. Our research also highlights a complex population structure, combining genomic genetic data at multiple scales and using spatial analysis of genetic structure to better understand genetic variation in widespread lineages or species. How environment (habitat)-mediated selection has influenced this hybrid zone remains an open question for future research.

Study limitations and future directions

Two questions remain unresolved in this study. First, *S. glandulatus* and *S. mammatus* are considered different species based on morphological characteristics (Fei et al., 2005) and both can be easily differentiated from *S. boulengeri*. Interestingly, based on the presence of mtDNA genetic markers, *S. glandulatus* and *S. mammatus* were nested in the *S. boulengeri* clades, as also reported by Li et al. (2009) and Chen et al. (2009). Here, intron genetic markers (cmyc-2) were successfully amplified, indicating that the two species diverged earlier from *S. boulengeri* than indicated by the mtDNA gene results, but exons were difficult to amplify, suggesting that hybridization may have only occurred at the early stage of divergence. Recent gene exchange may not have occurred due to resource partitioning and niche divergence, increasing the difficulty to understand the reasons for specific clade diversity and mechanisms of population evolution.

The second issue is that genetic analysis was based on a small number of genetic markers, which may have affected the accuracy and precision of the inferred patterns of genetic structure and demographic analyses. Therefore, our results should be treated with caution. For more complex models with additional populations and parameters, high-resolution genetic datasets may be necessary to obtain more precise and accurate estimates. These two caveats highlight the need for further genome sequencing and analysis.

CONCLUSIONS

We sequenced three mtDNA and four nuDNA genes, including one intron and three exons. We performed phylogenetic reconstructions based on 2 169 bp mtDNA fragments and explored population dynamics based on 3 286 bp nuDNA using a coalescent theory framework to infer divergence time, population size, and migration rates among three recently diverged clusters of *S. boulengeri*. The apparent mitonuclear discordance was likely the result of secondary contact hybridization, without significant sex-biased dispersal. We assessed the validity of these results by measuring the extent of admixture across secondary contact zones with cline model analysis. Assuming no significant geographic barriers to gene flow at the transition zones, we identified a wide hybrid zone around the Hengduan Mts. with no apparent molecular transition zone, suggesting that reproductive isolation was not established in the hybrid zone. Landscape genetic analysis of ecological drivers also suggested that the correlation between divergence in body size and Precipitation of driest month may have partially contributed to the establishment of IBE in this system. Our study provides a basis for evaluating population demographic changes in genetic structure and local adaptation in Xizang alpine toads with high-resolution genetic data, with several implications for conservation and invasion biology.

Our research also suggests that disregarding hybridization or gene flow will overestimate species delimitation (e.g., phylogenetic trees depending solely on mtDNA genes). Therefore, we caution against delimiting recently or rapidly diverging populations where gene flow may have been prevalent, even though environment-mediated selection may have structured this hybrid zone. For such cases, population genomics should be used in conjunction with population demographic models that account for gene flow to delimit cryptic species boundaries and evaluate lineage separation in line with the general lineage concept of species in the future.

SCIENTIFIC FIELD SURVEY PERMISSION INFORMATION

Permission for field surveys was issued by the Yunnan Forestry and Grassland Bureau (No. [2019]825), Forestry and Grassland Administration of Xizang Autonomous Region (No. [2021]71), and Chengdu Institute of Biology, Chinese Academy of Sciences ([2019]168). All specimen collection protocols were approved by the Animal Care Committee of the Chengdu Institute of Biology, Chinese Academy of Sciences.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

X.Q.L. and F.X. conceived the study; X.Q.L. and F.X. collected the data with assistance from Y.M.H., S.C.S., P.Y.Z., and J.P.J.; X.Q.L. performed the experiments and processed the raw data; X.Q.L., Y.M.H., and W.Z.Y. analyzed the data; X.Q.L., W.Z.Y., C.K.S., and F.X. wrote the manuscript. All co-authors contributed to edits and comments toward the final manuscript. All authors read and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

We thank Dr. Jinzhong Fu from the University of Guelph for suggestions on this manuscript, and three anonymous reviewers for their insightful comments on this study. We also thank Dr. Bin Wang for advice in exon amplification and Dr. Xiao-Xiao Shu and Miss Li-Ling Wang from Yunnan University for help provided to the first author in STRUCTURE and DIYABC analysis. We also thank Dr. Guo-Cheng Shu, Dr. Gang Wang, Dr. Bo Cai, Dr. Zhong-Yi Yao, Mr. Jin Zhou, and Mr. Cheng Shen for their kind help in field work. Special thanks are given to the Herpetological Museum of the Chengdu Institute of Biology, Chinese Academy of Sciences (CAS), and the Animal Branch of the Germplasm Bank of Wild Species, CAS, for providing specimens.

REFERENCES

Auguie B, Antonov A. 2017. Miscellaneous functions for "Grid" graphics (Version 2.3). https://cran.r-project.org/web/packages/gridExtra/gridExtra.pdf.

Banta J. 2020. How to convert a FASTA file to a STRUCTURE file [Video]. YouTube.https://www.youtube.com/watch?v=EO6AtZPgz1g.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**: 113−148.

Barton NH, Hewitt GM. 1989. Adaptation, speciation and hybrid zones. *Nature*, **341**(6242): 497−503.

Beerli P, Mashayekhi S, Sadeghi M, Khodaei M, Shaw K. 2019. Population genetic inference with MIGRATE. *Current Protocols in Bioinformatics*, **68**(1): e87.

Blanquart F, Gandon S, Nuismer SL. 2012. The effects of migration and drift on local adaptation to a heterogeneous environment. *Journal of Evolutionary Biology*, **25**(7): 1351−1363.

Boaratti AZ, Da Silva FR. 2015. Relationships between environmental gradients and geographic variation in the intraspecific body size of three species of frogs (Anura). *Austral Ecology*, **40**(8): 869−876.

Bohonak AJ. 1999. Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**(1): 21−45.

Bonnet E, Van de Peer Y. 2002. zt: a sofware tool for simple and partial mantel tests. *Journal of Statistical Software*, **7**(10): 1−12.

Broquet T, Petit EJ. 2009. Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics*, **40**: 193−216.

Bruen TC, Philippe H, Bryant D. 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics*, **172**(4): 2665−2681.

Brumfield RT, Jernigan RW, Mcdonald DB, Braun MJ. 2001. Evolutionary implications of divergent clines in an avian (*Manacus*: Aves) hybrid zone. *Evolution*, **55**(10): 2070−2087.

Bryant D, Moulton V. 2004. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, **21**(2): 255−265.

Buggs RJA. 2007. Empirical study of hybrid zone movement. *Heredity*, **99**(3): 301−312.

Chen W, Bi K, Fu JZ. 2009. Frequent mitochondrial gene introgression among high elevation Tibetan megophryid frogs revealed by conflicting gene genealogies. *Molecular Ecology*, **18**(13): 2856−2876.

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, et al. 2014. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**(8): 1187−1189.

Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ, et al. 2008. Inferring population history with *DIY ABC*: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**(23): 2713−2719.

Crease TJ, Lynch M, Spitze K. 1990. Hierarchical analysis of population genetic variation in mitochondrial and nuclear genes of *Daphnia pulex*. *Molecular Biology and Evolution*, **7**(5): 444–458.

Derryberry EP, Derryberry GE, Maley JM, Brumfield RT. 2014. HZAR: hybrid zone analysis using an R software package. *Molecular Ecology Resources*, **14**(3): 652−663.

Després L. 2019. One, two or more species? Mitonuclear discordance and species delimitation. *Molecular Ecology*, **28**(17): 3845−3847.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**(8): 1969−1973.

Dufresnes C, Pribille M, Alard B, Gonçalves H, Amat F, Crochet PA, et al. 2020. Integrating hybrid zone analyses in species delimitation: lessons from two anuran radiations of the Western Mediterranean. *Heredity*, **124**(3): 423−438.

Edwards S, Bensch S. 2009. Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Molecular Ecology*, **18**(14): 2930−2933.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**(8): 2611−2620.

Everson KM, Gray LN, Jones AG, Lawrence NM, Foley ME, Sovacool KL, et al. 2021. Geography is more important than life history in the recent diversification of the tiger salamander complex. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(17): e2014719118.

Excoffier L, Foll M, Petit RJ. 2009. Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, **40**: 481−501.

Fei L, Ye CY, Huang YZ, Jiang JP, Xie F. 2005. Chapter 2: taxonomic terms and metrics of amphibians. *In*: An Illustrated key to Chinese Amphibians. Sichuan Publishing House of Science and Technology, 3–5. (in Chinese)

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, **37**(12): 4302−4315.

Francis RM. 2017. POPHELPER: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, **17**(1): 27−32.

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, **34**: 397−423.

García-Girón J, García P, Fernández-Aláez M, Bécares E, Fernández-Aláez C. 2019. Bridging population genetics and the metacommunity perspective to unravel the biogeographic processes shaping genetic differentiation of *Myriophyllum alterniflorum* DC. *Scientific Reports*, **9**(1): 18097.

Han WX, Fang XM, Ye CC, Teng XH, Zhang T. 2014. Tibet forcing Quaternary stepwise enhancement of westerly jet and central Asian aridification: carbonate isotope records from deep drilling in the Qaidam salt

playa, NE Tibet. *Global and Planetary Change*, **116**: 68−75.
Head MJ, Gibbard PL. 2015. Early–Middle Pleistocene transitions: linking terrestrial and marine realms. *Quaternary International*, **389**: 7−46.
Hewitt GM. 1988. Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, **3**(7): 158−167.
Hijmans RJ, Karney C, Williams E, Vennes C. 2021. Geosphere: spherical trigonometry.https://rdrr.io/cran/geosphere/.
Hofmann S, Stöck M, Zheng YC, Ficetola FG, Li JT, Scheidt U, et al. 2017. Molecular phylogenies indicate a Paleo-Tibetan origin of himalayan lazy toads (*Scutiger*). *Scientific Reports*, **7**(1): 3308.
Hudson RR, Turelli M. 2003. Stochasticity overrules the "three-times rule": genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **57**(1): 182−190.
Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**(2): 254−267.
Huson DH, Bryant D. 2008. User manual for SplitsTree4 V4.10. http://nebc.nerc.ac.uk/bioinformatics/documentation/splitstree/manual.pdf.
Huson DH, Kloepper TH. 2005. Computing recombination networks from binary sequences. *Bioinformatics*, **21**(S2): ii159−ii165.
Jiang FC, Wu XH. 1998. Late cenozoic tectonic movement in geomorphologic boundary belt of southeastern Qinghai-Xizang Plateau. *Journal of Chengdu University of Technology*, **25**(2): 162−168. (in Chinese)
Jombart T, Collins C. 2015. A tutorial for discriminant analysis of principal components (DAPC) using *adegenet* 2.0. 0.https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf.
Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**: 94.
Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**(6): 587−589.
Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics Analysis across computing platforms. *Molecular Biology and Evolution*, **35**(6): 1547−1549.
Larson EL, White TA, Ross CL, Harrison RG. 2014. Gene flow and the maintenance of species boundaries. *Molecular Ecology*, **23**(7): 1668−1678.
Legendre P, Fortin MJ. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, **10**(5): 831−844.
Li R, Chen W, Tu LH, Fu JZ. 2009. Rivers as barriers for high elevation amphibians: a phylogeographic analysis of the alpine stream frog of the Hengduan Mountains. *Journal of Zoology*, **277**(4): 309−316.
Lin XQ, Shih CK, Hou YM, Shu XX, Zhang MH, Hu JH, et al. 2021. Climatic-niche evolution with key morphological innovations across clades within *Scutiger boulengeri* (Anura: Megophryidae). *Ecology and Evolution*, **11**(15): 10353−10368.
López-Uribe MM, Jha S, Soro A. 2019. A trait-based approach to predict population genetic structure in bees. *Molecular Ecology*, **28**(8): 1919−1929.
Major EI, Höhn M, Avanzi C, Fady B, Heer K, Opgenoorth L, et al. 2021. Fine-scale spatial genetic structure across the species range reflects recent colonization of high elevation habitats in silver fir (*Abies alba* Mill.). *Molecular Ecology*, **30**(20): 5247−5265.
Marko PB, Hart MW. 2011. The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution*, **26**(9): 448−456.
McQuillan MA, Rice AM. 2015. Differential effects of climate and species interactions on range limits at a hybrid zone: potential direct and indirect impacts of climate change. *Ecology and Evolution*, **5**(21): 5120−5137.
Medina R, Wogan GOU, Bi K, Termignoni-García F, Bernal MH, Jaramillo-Correa JP, et al. 2021. Phenotypic and genomic diversification with isolation by environment along elevational gradients in a neotropical treefrog. *Molecular Ecology*, **30**(16): 4062−4076.
Menon M, Bagley JC, Friedline CJ, Whipple AV, Schoettle AW, Leal-Sàenz A, et al. 2018. The role of hybridization during ecological divergence of southwestern white pine (*Pinus strobiformis*) and limber pine (*P. flexilis*). *Molecular Ecology*, **27**(5): 1245−1260.
Ottaviani D, Cairns SC, Oliverio M, Boitani L. 2006. Body mass as a predictive variable of home-range size among Italian mammals and birds. *Journal of Zoology*, **269**(3): 317−330.
Parins-Fukuchi C, Stull GW, Smith SA. 2021. Phylogenomic conflict coincides with rapid morphological innovation. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(19): e2023058118.
Pereira RJ, Singhal S. 2022. A lizard with two tales: what diversification within *Sceloporus occidentalis* teaches us about species formation. *Molecular Ecology*, **31**(2): 407−410.
Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155**(2): 945−959.
Prugnolle F, de Meeus T. 2002. Inferring sex-biased dispersal from population genetic tools: a review. *Heredity*, **88**(3): 161−165.
Rambaut A, Drummond A. 2012. FigTree v1.4. Molecular evolution, phylogenetics and epidemiology. http://tree.bio.ed.ac.uk/software/figtree/.
Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**(5): 901−904.
Sexton JP, Hangartner SB, Hoffmann AA. 2014. Genetic isolation by environment or distance: which pattern of gene flow is most common?. *Evolution*, **68**(1): 1−15.
Shen XX, Liang D, Feng YJ, Chen MY, Zhang P. 2013. A versatile and highly efficient toolkit including 102 nuclear markers for vertebrate phylogenomics, tested by resolving the higher level relationships of the caudata. *Molecular Biology and Evolution*, **30**(10): 2235−2248.
Shu XX, Hou YM, Cheng MY, Shu GC, Lin XQ, Wang B, et al. 2022. Rapid genetic divergence and mitonuclear discordance in the Taliang knobby newt (*Liangshantriton taliangensis*, Salamandridae, Caudata) and their driving forces. *Zoological Research*, **43**(1): 129−146.
Song ZM, Huang DM, Chang C. 1990. On the development and population age structure of *Scutiger boulengeri* tadpoles. *Acta Zoologica Sinica*, **36**(2): 187−193. (in Chinese)
Stevens VM, Trochet A, Blanchet S, Moulherat S, Clobert J, Baguette M. 2013. Dispersal syndromes and the use of life-histories to predict dispersal. *Evolutionary Applications*, **6**(4): 630−642.
Szymura JM, Barton NH. 1991. The genetic structure of the hybrid zone between the fire-bellied toads *Bombina Bombina* and *B. Variegata*: comparisons between transects and between loci. *Evolution*, **45**(2): 237−261.
Taylor SA, Larson EL, Harrison RG. 2015. Hybrid zones: windows on climate change. *Trends in Ecology & Evolution*, **30**(7): 398−406.
Toews DPL, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**(16): 3907−3930.
Tsuda Y, Nakao K, Ide Y, Tsumura Y. 2015. The population demography of *Betula maximowicziana*, a cool-temperate tree species in Japan, in relation to the last glacial period: its admixture-like genetic structure is the result of

simple population splitting not admixing. *Molecular Ecology*, **24**(7): 1403−1418.

Twyford AD, Wong ELY, Friedman J. 2020. Multi-level patterns of genetic structure and isolation by distance in the widespread plant *Mimulus guttatus*. *Heredity*, **125**(4): 227–239.

Urquhart J, Wang YZ, Fu JZ. 2009. Historical vicariance and male-mediated gene flow in the toad-headed lizards *Phrynocephalus przewalskii*. *Molecular Ecology*, **18**(17): 3714–3729.

Valenzuela-Sánchez A, Cunningham AA, Soto-Azat C. 2015. Geographic body size variation in ectotherms: effects of seasonality on an anuran from the southern temperate forest. *Frontiers in Zoology*, **12**: 37.

Wang IJ. 2013. Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, **67**(12): 3403−3411.

Wang K, Lenstra JA, Liu L, Hu QJ, Ma T, Qiu Q, et al. 2018. Incomplete lineage sorting rather than hybridization explains the inconsistent phylogeny of the wisent. *Communications Biology*, **1**(1): 169.

Wen GN, Fu JZ. 2021. Isolation and reconnection: demographic history and multiple contact zones of the green odorous frog (*Odorrana margaretae*) around the Sichuan Basin. *Molecular Ecology*, **30**(16): 4103−4117.

Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**(3): 1177−1191.

Xing YW, Ree RH. 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(17): E3444−E3451.

Yang WZ, Feiner N, Pinho C, While GM, Kaliontzopoulou A, Harris DJ, et al. 2021. Extensive introgression and mosaic genomes of Mediterranean endemic lizards. *Nature Communications*, **12**(1): 2762.

Zhang D, Gao FL, Jakovlić I, Zou H, Zhang J, Li WX, et al. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources*, **20**(1): 348−355.