



DNA Barcodes and Species Distribution Models Evaluate Threats of Global Climate Changes to Genetic Diversity: A Case Study from *Nanorana parkeri* (Anura: Dicroglossidae)

Wei-wei Zhou¹, Bao-lin Zhang^{1,2}, Hong-man Chen¹, Jie-qiong Jin¹, Jun-xiao Yang¹, Yun-yu Wang¹, Ke Jiang¹, Robert W. Murphy^{1,3}, Ya-ping Zhang^{2*}, Jing Che^{1*}

1 State Key Laboratory of Genetic Resources and Evolution, and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China, **2** Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming, China, **3** Centre for Biodiversity and Conservation Biology, Royal Ontario Museum, Toronto, Ontario, Canada

Abstract

Anthropogenic global climate changes are one of the greatest threats to biodiversity. Distribution modeling can predict the effects of climate changes and potentially their effects on genetic diversity. DNA barcoding quickly identifies patterns of genetic diversity. As a case study, we use DNA barcodes and distribution models to predict threats under climate changes in the frog *Nanorana parkeri*, which is endemic to the Qinghai-Tibetan Plateau. Barcoding identifies major lineages W and E. Lineage W has a single origin in a refugium and Lineage E derives from three refugia. All refugia locate in river valleys and each greatly contributes to the current level of intraspecific genetic diversity. Species distribution models suggest that global climate changes will greatly influence *N. parkeri*, especially in the level of genetic diversity, because two former refugia will fail to provide suitable habitat. Our pipeline provides a novel application of DNA barcoding and has important implications for the conservation of biodiversity in southern areas of the Qinghai-Tibetan Plateau.

Citation: Zhou W-w, Zhang B-l, Chen H-m, Jin J-q, Yang J-x, et al. (2014) DNA Barcodes and Species Distribution Models Evaluate Threats of Global Climate Changes to Genetic Diversity: A Case Study from *Nanorana parkeri* (Anura: Dicroglossidae). PLoS ONE 9(8): e103899. doi:10.1371/journal.pone.0103899

Editor: Dirk Steinke, Biodiversity Institute of Ontario - University of Guelph, Canada

Received: January 19, 2014; **Accepted:** July 7, 2014; **Published:** August 5, 2014

Copyright: © 2014 Zhou et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by grants of the Strategic Priority Research Program (B) (XDB03030113), the Ministry of Science and Technology of China (MOST) (2014FY210200), and the “Western Light” Talents Training Program of the Chinese Academy of Sciences (CAS) to ZWW; the Key Research Program (KJZD-EW-L07) of CAS, the National Natural Science Foundation of China (31090250), MOST (2011FY120200), and the Bureau of Science and Technology of Yunnan, China (2010CI045) to JC; and the Major Innovation Program of CAS (KSCX2-EW-Z-2) to ZYP. Research also was supported by Discovery Grant 3148 from the Natural Sciences and Engineering Research Council of Canada to RWM. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: chej@mail.kiz.ac.cn (JC); zhangyp@mail.kiz.ac.cn (YPZ)

Introduction

Climatic changes influence organisms and an understanding of how this occurs is important for conservation. More than one line of evidence documents the impact anthropogenic global climate change (GCC) exerts on organisms [1]. Explorations into how past climate changes influenced organisms may serve to predict future impacts of GCC. Genetic diversity is important in conservation because higher levels maintain the evolutionary potential of species. However, the distribution of genetic diversity is often uneven across the range of a species and many factors may contribute to this. Environmental changes during glacial-interglacial cycling in the Quaternary is one of the most important historical drivers of genetic patterns [2,3]. For example, by retaining suitable habitat over several glacial cycles, refugia hold higher levels of genetic diversity compared with recently occupied areas [2,3]. Refugia also drive genetic distinctiveness within species owing to providing long-term geographic isolation. A clear understanding a species' evolutionary history and its drivers is important for planning conservation [4].

Genetic analyses form the cornerstone of conservation planning, especially in defining objective evolutionarily significant units (ESUs) and management units (MUs) [5]. DNA barcoding [6], which uses a short, universal genetic marker (*COI* in eukaryotes) to identify matrilineal and species, may serve to efficiently identify ESUs and MUs.

Species distribution models (SDMs) also provide information useful for conservation planning [7–9]. Comparisons between SDMs for current and the Last Glacial Maximum (LGM) may provide a chance to explore the impacts of past climate changes to organisms. As an extension of this application, SDMs can compare the distributions of current and future habitats. This allows assessments of the risk of local extirpation and extinction caused by future habitat degradation. Whereas both DNA barcoding and SDMs provide valuable insights, their synthesis may serve to evaluate threats of GCC to organisms.

The Qinghai-Tibetan Plateau (QTP), which covers more than 2.5 million km² and has an average elevation of about 4000 m above sea level [10], is the largest and highest plateau on Earth. Unlike North America and Europe, no unified ice sheet formed on the QTP during the LGM, yet its environment changed

substantially. The average temperature was from 6°C to 9°C lower than today and precipitation decreased by from 30% to 70% [11].

In addition to climatic drivers, geographic features also play a role in the formation of patterns of genetic diversity. In the southern QTP, two mountains stretch from east to west and the Yarlung Zangpo River (YZR; Brahmaputra River) (Fig. S2) occurs between them. The Nianqingtangula Mountains (NM; Nyainqentanglha Mountains) in the east and Gangdisi Mountains (GM; Kailas Range) in the west form the northern mountains and the Himalayas form the southern boundary. Several rivers flow southward across the latter. The complex geography may have offered refugia during dramatic Quaternary climatic changes. Analyses of matrilineal dispersal are likely to reveal such history. Driven by the GCC, future severe temperature changes of the QTP may exceed those of lower elevations [12]. Thus, the GCC may impose enormous impacts on organisms living in the QTP. We assess this possibility by synthesizing DNA barcoding and SDMs analyses.

Because of their relatively low mobility and physiological requirements, amphibians are sensitive to environmental changes, especially temperature and precipitation. Further, amphibians retain high-resolution signals of historical responses to environmental perturbations [13]. The environment on the QTP is harsh for amphibians; suitable environment is rare except in the eastern and southern edges. A previous study on *Rana kukunoris*, a common frog on eastern QTP, revealed a unique genetic pattern compared to other taxa in the region [14]. Comparative amphibian studies from the southern QTP will yield insights into how environmental changes affect the biota fauna of the QTP.

Nanorana parkeri, a median-sized frog belonging to the family Dicroglossidae, is endemic to southern and southeastern Tibet. It occurs at elevations ranging from 2800 to 5000 m in valleys of the YZR and north of the NM [15]. Climate changes during LGM greatly affected this area. Our study evaluates samples from across the entire range of *N. parkeri*. Using the species as model system, we test the effectiveness of a pipeline that combines DNA barcoding and SDMs to identify future threats to genetic diversity in the face of GCC.

Materials and Methods

Ethics Statement

All the species included in our study (Table S1) are not endangered or protected species according to the “Law of the People’s Republic of China on the Protection of Wildlife” and “Regulations for the Implementation of the People’s Republic of China on the Protection of terrestrial Wildlife” (State Council Decree [1992] No. 13). The permission for field work in four major management areas (39 localities) including Xigaze, Lhasa, Nyingchi and Qamdo was issued by the Forestry Department of Xizang [Tibet] Autonomous Region, China. Local people of the Forestry Bureaus involved and helped during the whole survey. Totally, 549 tissue samples including toe tips, muscle, livers, tadpoles and egg masses were obtained, following the Animal Use Protocols approved by the Animal Care and Ethics Committee of the Kunming Institute of Zoology, Chinese Academy of Sciences. For egg masses, we separated 5–10 eggs from egg mass and stored them in 95% alcohol after removing egg jelly. For adult, within each locality only five individuals were euthanized using clove oil firstly. Following euthanization, tissues dissected from adult specimens were preserved in 95% ethanol. More other adults were just cut two toe tips and then released. Tadpoles were also euthanized using clove oil firstly, then stored in 95% alcohol after removing gut. Table S1 lists our samplings information, including

species name, locality, GPS coordinates, and accession nos. in Genbank etc.

Population sampling and laboratory techniques

Our samplings covered the entire documented distribution range of *N. parkeri* (Fig. 1, Table S1). One individual of *N. pleskei*, *N. ventripunctata* and *N. liebigii* was used as an outgroup representative [16,17].

We extracted genomic DNA using standard phenol–chloroform extraction protocol [18]. Partial sequences of cytochrome c oxidase subunit I (*COI*) were sequenced for all individuals using universal primers [19]. PCR products were purified and used as the template DNA for cycle sequencing reactions performed using BigDye Terminator Cycle Sequencing Kit (v.2.0, Applied Biosystems, Foster City, USA) and an ABI PRISM 3730 DNA Analyzer.

Sequence alignment and phylogenetic analyses

Nucleotide sequences were checked by eye using LASERGENE 7.0 and aligned using CLUSTALX 1.81 [20] with default parameters. Subsequently, the aligned sequences were checked and optimized in MEGA 4.0 [21]. Identical haplotypes for mtDNA were collapsed using DNASP 5.10 [22]. The overall value of nucleotide diversity (π) and haplotype diversity (H) were also estimated using DNASP.

Phylogenetic analyses of the *COI* data were conducted using Bayesian inference (BI), maximum likelihood (ML) and maximum parsimony (MP). BI analyses were performed using MrBayes 3.1.2 [23]. We tested three different partition strategies based on codon positions (no partition; 1+2, 3; and 1, 2, 3). The best strategy was chosen based on the Bayes factor test [24] (Table 1), as it represented a robust method testing partitioning strategies [25]. Nucleotide substitution models were selected for each data partition using the Akaike information criterion in MrModeltest v2.3 [26]. BI analyses for each partition strategy were run 3 million generations while sampling trees every 1000 generations. The first 50% of the sampled trees were discarded as burn-in. The final analyses employing the best partition strategy were run for 10 million generations. AWTY [27] was used to confirm satisfactory convergence of the topological split frequencies. MP analyses were implemented using PAUP* 4.0b10 [28]. Heuristic searches with TBR were executed for 1000 random pseudoreplicates with all characters treated as unordered and equally weighted. Bootstrap analyses were conducted using 1000 replicates to assess nodal reliabilities. ML analyses were conducted using RaxML 7.0.4 [29] and the GTR+I+G model was implemented for each data partition.

We built a median-joining network (MJN) using NETWORK 4.5 [30] to visualize the frequencies of the haplotypes. To remove excessive links and median vectors, we used the MP option [31].

Divergence time estimation

Different demographic and molecular clock models were compared using path sampling and stepping-stone sampling (Table 2) [32]. Employing the best models, time to most recent common ancestor was estimated using BEAST 1.7.5 [33]. Due to the absence of a reliable fossil record and an established substitution rate of *COI* for *N. parkeri*, we used the divergence time between *N. parkeri* and *N. pleskei* (8.9 ± 2.7 Ma) [16] as a secondary calibration point. Model selection and partition scheme were the same as used in the BI analyses. The final analyses involved two independent runs of 30 million generations each, while sampling trees every 1000th generation. Burn-in and convergence of the chains were determined with Tracer 1.5

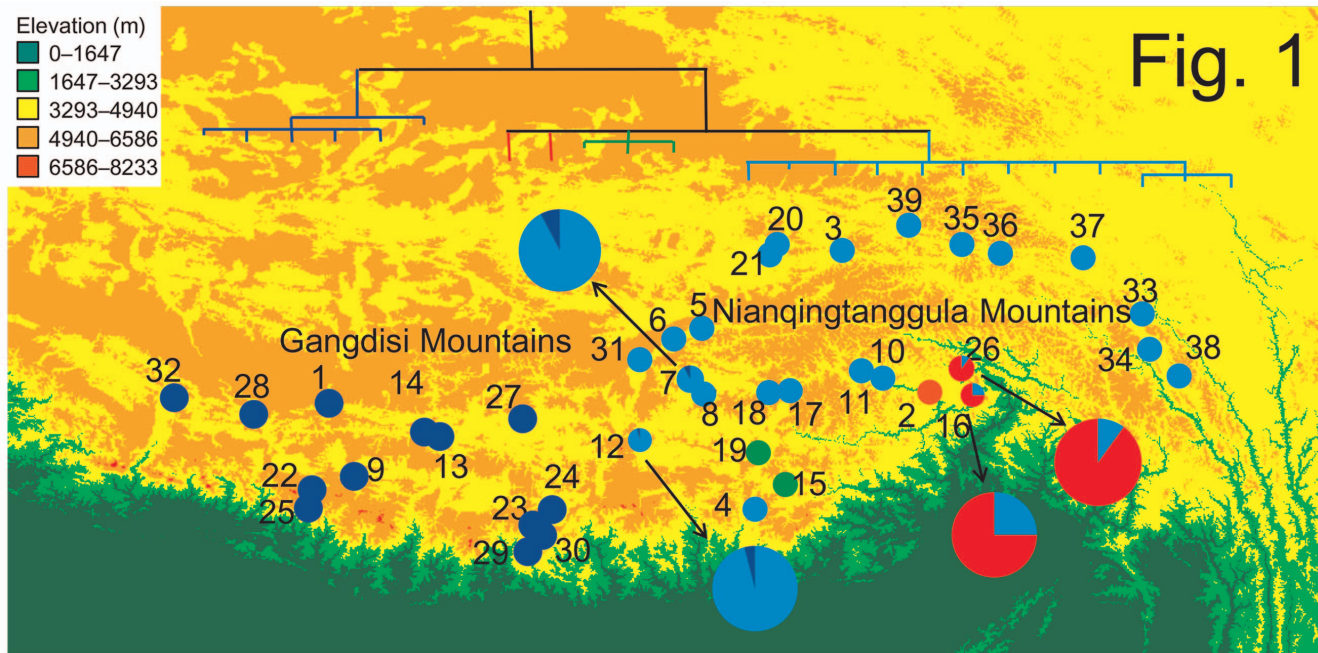


Figure 1. Sampling sites and the maternal genealogy based on mtDNA sequence data. Site numbers refer to Table S1. Colors of the matrilineal lines match those of samplings localities on the map. Four localities containing two matrilineal lineages each are shown in larger pie figures. doi:10.1371/journal.pone.0103899.g001

[34]. The measures of effective sample sizes were used to determine the Bayesian statistical significance of each parameter.

Population structure and demographic analyses

We explored population structure and genetic diversity landscape using SPADS 1.0 [35]. Groups of populations were defined as for SAMOVA [36]. We explored values of K (number of groups) ranging from 2 to 10 with 100 simulated annealing processes. The optimum value of K was identified by exploring the behavior of the indices F_{CT} and F_{SC} . Spatial patterns of genetic diversity based on allelic richness (Ar) and π across the landscape were explored using the GDivPAL function in SPADS.

To test for the influences of mountains and rivers on the genetic structure of *N. parkeri*, we measured the population structure by three independent analyses of molecular variance (AMOVA) [37]: populations north and south of the YZR; populations north of NM, west of the boundary of GM and NM, and the remaining ones; and four groups according to the results of the SAMOVA. Analyses were performed using Arlequin 3.5 [38] and significance was assessed by 10000 permutations.

We investigated past changes of effective population size using the neutral test and mismatch distributions. Tajima's D [39] and Fu's F_s statistics [40] with 10000 coalescent simulations were calculated using Arlequin. Pairwise mismatch distributions [41]

were calculated for each group in Arlequin. The expected distribution under a model of sudden demographical expansion was generated using 10000 permutations. The significance of deviations from this model was tested using the sum of squared deviation (SSD) and raggedness index (Rag). All analyses were performed for each lineage separately because population subdivisions could have masked the effects of expansions.

Species distribution modeling

We inferred the potential geographic range of *N. parkeri* using the maximum entropy model implemented in MAXENT 3.3.3 [42,43]. Environmental variables from the WorldClim database with resolutions of 2.5 arc-minutes [44] were downloaded as environmental layers. Because of the controversy about whether correlated variables should be removed or not, we used all of the 19 bioclimatic layers [45]. All layers were cropped to span from 83°E to 99°E and from 26°N to 33°N.

Random null distributions were built to test for the significance of our SDM. For this test, we built a new SDM using 39 random points (as same number as our sampling localities). This process was repeated 100 times and the areas under the curves (AUCs) were used as null distributions [46].

Assuming niche conservatism over time [47–49], we predicted the former distribution of *N. parkeri* by projecting our model on

Table 1. Bayes factor of each partition strategy.

Partition Strategy	Ln(BF)	Standard Deviation
no partition	-1885.375	0.134
1+2, 3	-1844.969	0.14
1, 2, 3	-1885.375	0.12

$2\ln BF (H_1 - H_0) > 10$ was treated as decisive support for each hypothesis.

doi:10.1371/journal.pone.0103899.t001

Table 2. Results of demographic model selection.

	strick		lognormal relaxed clock	
	PS	SS	PS	SS
constant	-1553.29	-1553.31	-1553.33	-1553.34
exponential	-1556.44	-1556.53	-1556.24	-1556.27
logistic	-1557.29	-1557.36	-1555.73	-1555.77
expansion	-1554.42	-1554.45	-1555.05	-1555.07
bayesian skyline	-1544.51	-1544.56	-1545.42	-1545.46

PS: path sampling; SS: stepping-stone sampling.
doi:10.1371/journal.pone.0103899.t002

LGM climatic layers. Predicted distributions during the LGM were generated by downloading both the community climate system model (CCSM) [50] and model for interdisciplinary research on climate (MIROC) [51] from the WorldClim database. To predict the influences of GCC to this species, we also projected the model to climate data of the 2080 s based on MIROC model under the A1b scenario. In this way, we predicted the distribution changes in future.

Results

Sequence information

We obtained 549 partial *COI* sequences from 39 localities of *N. parkeri*, plus one sequence each from *N. pleskei*, *N. ventripunctata* and *N. liebigii*. The fragment consisted of 539 base pairs, of which 35 positions exhibited variation and 27 were potentially parsimony informative, resulting in 23 haplotypes in *N. parkeri*. All sequences were deposited in GenBank (Table S1). The overall value of π was 0.01746 ± 0.00055 and H was 0.678 ± 0.016 .

Genealogical analyses of *COI*

Bayes factor test showed a preference for the partition strategy of 1+2, 3 (Table 1). The best fit substitution model for the first and second codon was HKY and GTR+G model was the best fit for third codon.

Matrilineal genealogies obtained from BI, ML and MP analyses were nearly identical (Fig. 1 & Fig. S1) and all methods recovered lineages East (E) and West (W). The boundary of GM and NM separated these two lineages. Lineage W was comprised of 6 haplotypes, which had no clear structure. Lineage E was comprised 17 haplotypes of which 12 formed highly supported sublineage E1. The MJN depicted patterns similar to those of the gene tree (Fig. 2). Haplotypes H1 and H11, which were more common than the other haplotypes, occupied central positions.

Divergence time estimates

Results of model comparisons were shown in Table 2. Both path sampling and stepping-stone sampling suggested that the BSP demographic model under strict clock outperformed the alternative models. The average divergence time estimations were shown in Fig. 3. Lineages E and W diverged about 1.4 Ma (95% CI: 0.6–2.4 Ma). Lineage E diverged about 0.41 Ma (95% CI: 0.15–0.81 Ma) and the radiation of Lineage W happened about 0.15 Ma (95% CI: 0.05–0.32 Ma).

Genetic structure and demographic history

Results of population clustering were illustrated in Figure 4. F_{CT} plateaued when K was 4 (Table S2). Samples from Lineage W

grouped together (Fig. 4, dark blue). Lineage E contained 3 groups, of which populations in the southern edges of the QTP formed two small groups (localities 2, 16 and 26; and locality 15 and 19); the remaining localities comprised the third group (Fig. 4, light blue). Spatial patterns of genetic diversity based on A_r and π were shown in Figure 5 and listed in Table S3. Several peaks indicated high levels of genetic diversity across the distribution of *N. parkeri*. This was congruent with the pattern of lineage divergence.

When forming two groups based on the YZR, among-group diversity accounted for 16.04% of the overall variation and among-populations within groups accounted for 80.76%. Dividing populations according to the mountains, among-group diversity accounted for 87.90% of the overall variation and among-populations within

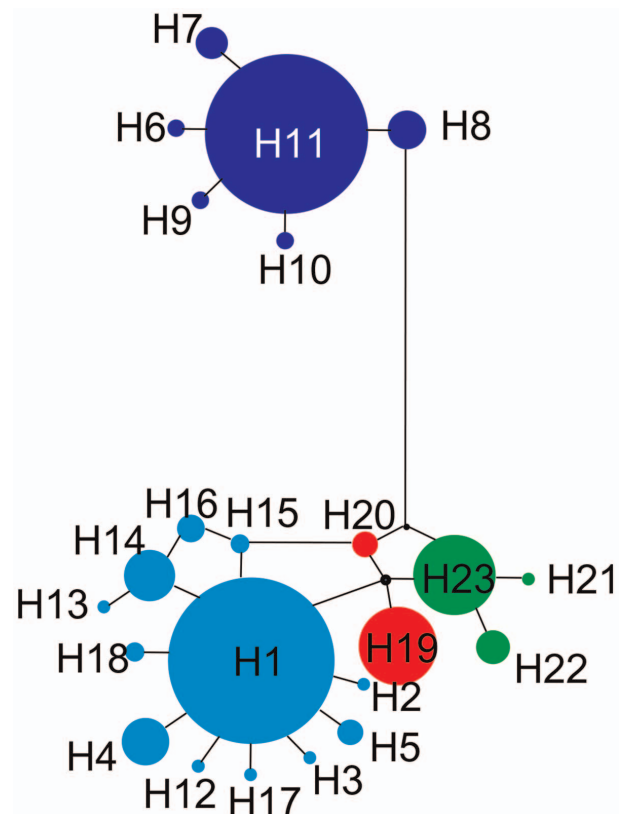


Figure 2. The median-joining network of *COI* haplotypes. Colors correspond to maternal lineages in Figure 1.
doi:10.1371/journal.pone.0103899.g002

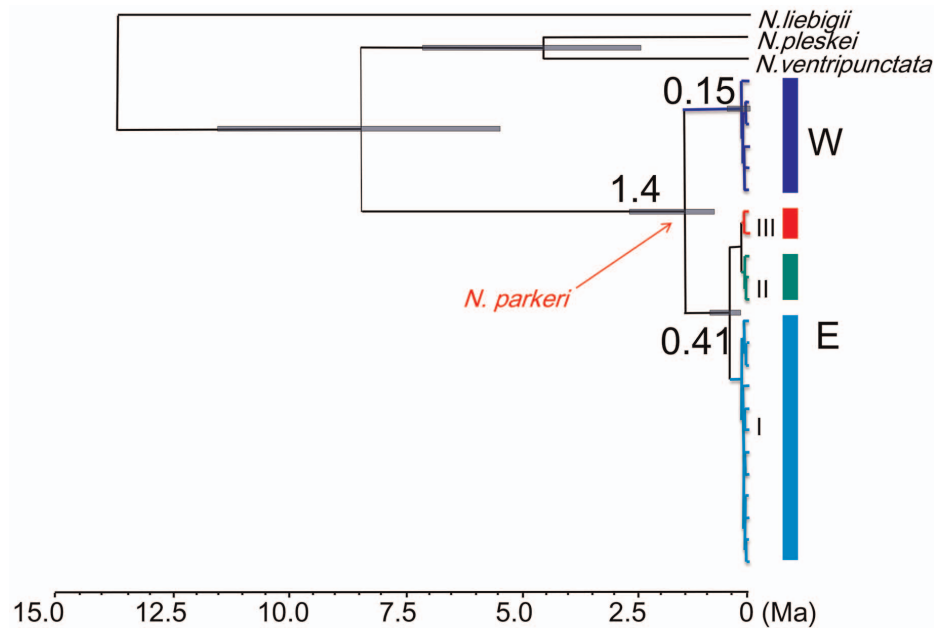


Figure 3. Estimates of divergence times obtained using BEAST 1.7.5.
doi:10.1371/journal.pone.0103899.g003

groups accounted for 9.67%. Finally, after dividing populations based on results of the SAMOVA, among-group diversity accounted for 96.88% of the overall variation and among-populations within groups accounted for 0.97% (Table 3).

For Lineage W, the values of Tajima's D and Fu's F_s were significantly negative (Table 4), which indicated population expansions. The hypothesis of sudden expansion was not rejected by mismatch distribution analyses (Fig. 6) as the SSD and Rag were insignificant ($P_{SSD} = 0.381$ and $P_{RAG} = 0.658$). In sublineage E1, significantly negative values of both Tajima's D and Fu's F_s supported population expansions. The mismatch distribution analyses did not reject the sudden expansion model ($P_{SSD} = 0.545$ and $P_{RAG} = 0.609$). In sublineages E2, Tajima's D and Fu's F_s were negative, but insignificant. In sublineage E3, Tajima's D was negative but Fu's F_s was positive, yet all values were not significant. The null hypothesis of sudden expansion model was not rejected by mismatch distribution analyses.

Species distribution modeling

The AUCs from random data ranged from 0.605 to 0.800 (0.733 ± 0.033 ; Table S4). For the SDMs of *N. parkeri*, the AUC of our data was 0.904 and this was significantly better than that of a random model ($P < 0.001$).

The MIROC model predicted that the distribution of *N. parkeri* at the LGM was much smaller than that of today (Fig. 7). The results based on CCSM model suggested a similar pattern, but with larger areas than those from MIROC model. Both models suggested range fluctuations after the LGM in northern, high altitude areas. The boundary area between GM and NM, where lineages W and E divided, has been unsuitable for *N. parkeri* since the LGM. The predicted distribution under GCC assuming the A1b scenario was shown in Figure 7. *Nanorana parkeri* is predicted to have a larger distribution in the northwestern QTP. However, the species' distribution will contract in the southeastern QTP.

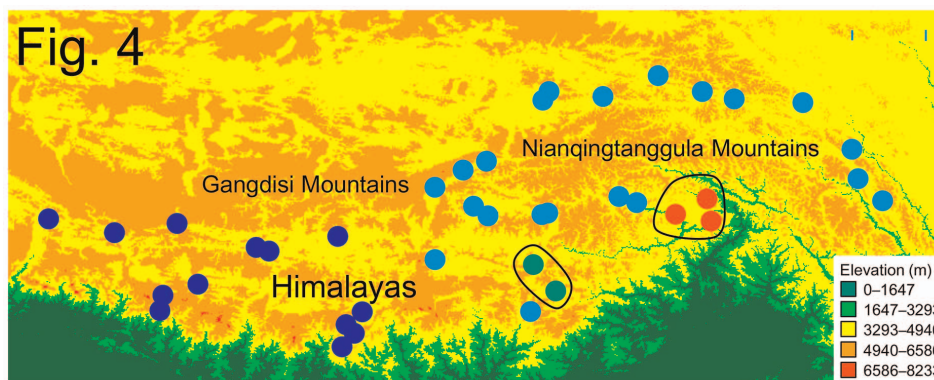


Figure 4. Results of SAMOVA. Different colors represent different groups. A solid line encircles two groups of Lineage E from microrefugia.
doi:10.1371/journal.pone.0103899.g004

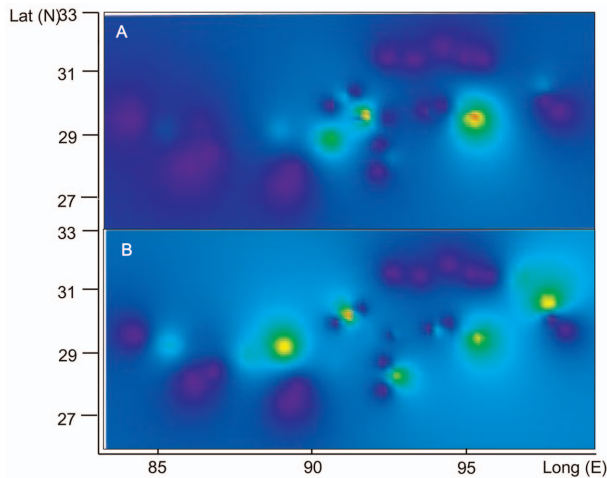


Figure 5. Genetic diversity across the range of *Nanorana parkeri*. X axis represents longitude and the Y axis latitude. A: π , pattern of nucleotide diversity; B: Ar, pattern of allelic richness. doi:10.1371/journal.pone.0103899.g005

Discussion

Lineage divergence and ESUs

ESUs, which can be determined by genetic differentiation at neutral markers caused by isolation, are important when considering conservation actions. Our study builds on the framework for delineating ESUs of *N. parkeri* by reconstructing the patterns of lineage divergence and its drivers. Our analyses suggest the recognition of ESUs that correspond to lineages W, E1, E2 and E3. Our research demonstrates that *COI* can quickly delineate ESUs.

The gene trees and MJN depict a clear east-west split for *N. parkeri*. SDMs suggest that unsuitable habitat between the GM and NM, both now and during the LGM, drives this pattern. The boundary is the demarcation line of two QTP climatic zones that are largely congruent with the 400 mm precipitation line. East of this line, the region is semi-humid, and to the west semi-arid [52]. Similar pattern was also found in *Hippophae tibetana* [53,54].

Geographic features, such as rivers or mountains, appear to contribute little to the patterns of genetic divergence. The YZR, which would impose a north-south split, does not drive population divergence. The AMOVA explains only 16.04% of the genetic variation in this scenario. Elevations of NM range from 5000 m to 6000 m, but no substantial genetic divergence distinguishes frogs in south and north of NM. Lineages E and W are separated by the boundary of GM and NM. However, habitat barriers and environmental differences more likely generate this pattern than geographical features. Two observations support this hypothesis. First, populations of lineages W and E both occur south of the GM

and NM. Second, *N. parkeri* occurs along the YZR valley, which flows across the boundary and connects lineages W and E.

The pattern of genetic divergence within each major lineage differs. The 6 haplotypes in Lineage W do not yield a clear pattern (Fig. 1). The network depicts a star-like haplogroup (Fig. 2). SAMOVA also suggests populations from W comprise a single group. In contrast, Lineage E differs in having three groups (Figs. 1, 2). Results from SAMOVA present the same pattern. Most northern populations form a group as do frogs from localities 15, 19 (E2) and localities 2, 16, 26 (E3). The difference in patterns between E and W may owe to topography. The intricate topography of the East Himalayas, which is a global biodiversity hotspot, provides suitable habitats and generates barriers that limit dispersal for Lineage E. Other species exhibit a similar pattern; populations in the fringes or southern slope of Himalayas are genetically different from those on the QTP [54,55]. Expansion and contraction of distribution ranges caused by climate changes provided opportunity for populations divergence in QTP [54–57].

The genetic pattern within *N. parkeri* suggests recognition of four ESUs, W, E1, E2 and E3. Certainly gene flow occurs within and between lineages E and W, as lineages or sublineages mix at localities 8, 12, 16 and 26. However, matrilineal uniqueness occurs in populations far from the boundary.

Glacial refugia and genetic diversity of *N. parkeri*

Refugia in the QTP prevented extinction during the LGM. In doing so, they conserved high levels of genetic diversity. In contrast, genetic diversity in recently occupied areas is much lower because of founder effects [2,3]. The identification of refugia is an important part of conservation because these areas preserve genetic diversity. Synthesizing DNA barcoding and SDMs, we quickly and effectively identify refugia for *N. parkeri*. Populations in refugia retain higher levels of diversity than population in newly occupied places (Fig. 5 and Table S3).

The multiple refugia hypothesis corresponds with patterns of mtDNA divergence. The matrilineal genealogy, MJN and SAMOVA analyses indicate the absence of genetic structure within Lineage W. Thus, Lineage W likely originates from a single refugium. The SDMs indicate the presence of suitable habitat in the river valley near locality 27 (Fig. 7) and the area retains a much higher level of genetic diversity (Fig. 5) than other sites. The neutral test and mismatch distribution analyses clearly detect a population expansion in Lineage W.

Our analyses identify several refugia in Lineage E. Sublineage E1 may originate from a northern refugium near localities 7 and 8, which occur in the valley at the confluence of the Lhasa River and the YZR (Fig. 7). During the LGM the suitable environment harbored a high level of genetic diversity. The absence of genealogical structure in sublineage E1 is congruent with a sudden population expansion, as are significantly negative values of Tajima's *D* and Fu's *F_s*, and the mismatch distribution analyses. Further, our analyses identify two microrefugia in the river valleys along southern edges of the QTP. Localities 2, 16 and 26, which

Table 3. Results of AMOVA.

Group compositions	Among group	Among populations	Within populations	Φ_{SC}	Φ_{ST}	Φ_{CT}
By River	16.04%	80.76%	3.20%	0.96183	0.96795	0.16039
By Mountain	87.90%	9.67%	2.42%	0.79973	0.97578	0.87905
By SAMOVA	96.88%	0.97%	2.16%	0.30944	0.97844	0.96877

doi:10.1371/journal.pone.0103899.t003

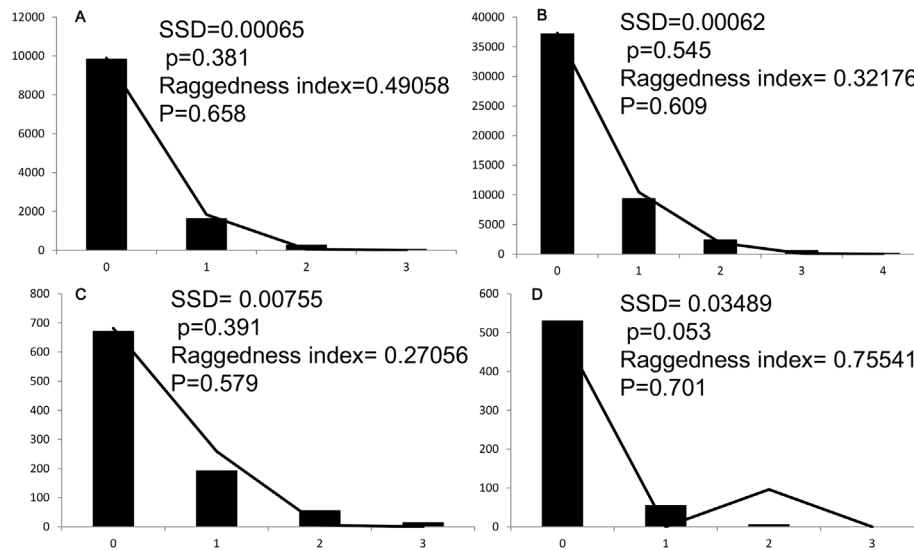


Figure 6. Results of the mismatch distribution for *Nanorana parkeri*. Bars indicate simulated mismatch distributions and lines denote the observed mismatch distributions. A: lineage W; B: sublineage E1; C: sublineage E2; D: sublineage E3. doi:10.1371/journal.pone.0103899.g006

are in the YZR Valley near Nyingchi and Medog, seemingly constitute a microrefugium. The other microrefugium consists of localities 15 and 19, which are also in a river valley. This river flows southwards across the Himalayas. The lower elevations of the river valleys make them less susceptible to the influences montane glaciers. River valleys in the Himalayas, especially those connected with southern slopes, appear to offer microrefugia.

Populations from sampling localities north of the NM also retain a high level of genetic diversity (Fig. 5). Private haplotypes H4 (locality 38), H2 and H3 (locality 33), and H5 (locality 37) occur here (Table S1). The SDMs suggest that suitable habitat existed in northern areas of NM. This area appears to harbor another microrefugium. However, these private haplotypes fail to cluster together and the most common haplotype in the area is H1 (Fig. 2 & Table S1). Thus, we cannot reject the hypothesis that these private haplotypes originated during population expansions after the LGM.

Potential threats of GCC to *N. parkeri*

Our study suggests that the combination of DNA barcoding and SDMs can detect threats of GCC to the survival of species. This pipeline facilitates conservation. SDMs predict that suitable habitat for *N. parkeri* may experience great shifts in the near future. Whereas the Northwest QTP will offer suitable habitat for *N. parkeri*, the Southeast QTP will become unsuitable. Although suitable habitat will experience an overall expansion, populations in the Southeast QTP may experience sharp decreases in

population size or become extirpated. Given that amphibians have poor dispersal abilities, the latter scenario may be an unfortunate consequence of GCC.

Our pipeline suggests that many populations of Lineage E may suffer from developing unsuitable habitats. The microrefugium near Nyingchi and Medog will become unsuitable under this prediction (Fig. 7). Fortunately, suitable habitat may persist in the two other refugia. Thus, genetic variation will decrease yet all may not be lost. For Lineage W, suitable habitat will disappear near the refugium at locality 27. Loss of this refugium will greatly decrease the amount of genetic diversity.

Our analyses suggest that genetic diversity of *N. parkeri* may greatly decrease. Although this facilitates effective conservation planning, we urge caution. Our SDMs for the future assume A1b scenarios, which involve rapid global economic growth with a balance of fossil and non-fossil energy sources. Each of possible scenarios yields different predictions about the future climate. All scenarios require constant adjustment according to global economic conditions. The complex landscape in the southern QTP may supply suitable microhabitats for *N. parkeri* which cannot be detected by our SDMs. Accordingly, we suggest the urgent development of an effective monitoring program, especially for populations in refugia that may lose suitable habitats.

Conclusion

DNA barcoding detects the genetic structure of *N. parkeri* and serves to define ESUs. Our analyses recover major lineages E and

Table 4. Statistics of neutrality tests for each lineage and sublineage.

	Tajima's D	P value	Fu's Fs	P value
W	-1.62115	0.0044	-6.41652	0.0005
E1	-1.80395	0.0033	-14.10103	0
E2	-0.69425	0.2176	-0.77698	0.1858
E3	-0.65715	0.2279	0.92064	0.5177

doi:10.1371/journal.pone.0103899.t004

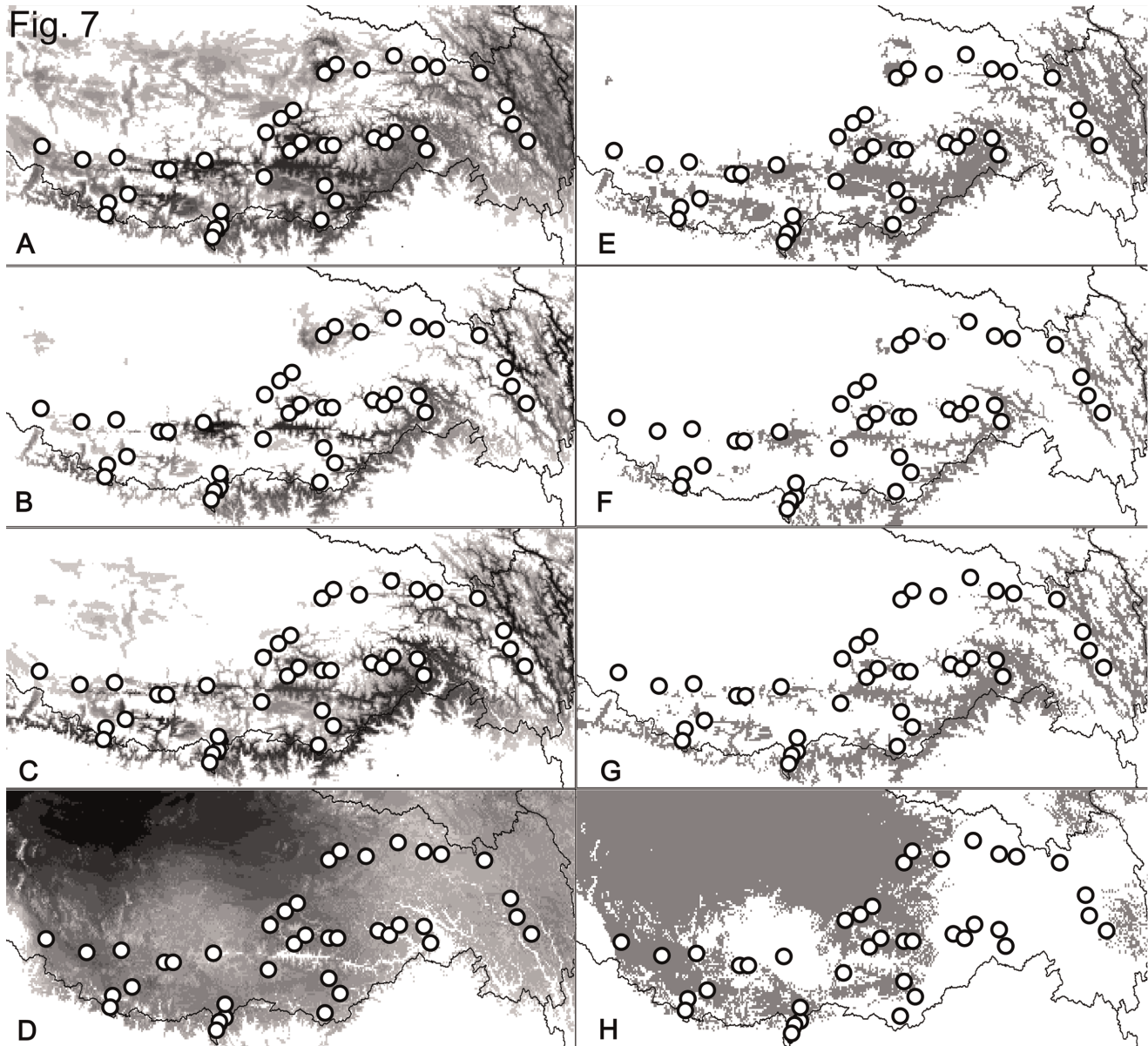


Figure 7. The predicted distribution of *Nanorana parkeri*. A: predicted distribution based on current data; B: distribution during the LGM based on MIROC model; C: distribution during the LGM based on CCSM model; D: the predicted distribution in the 2080 s under GCC; E–H: predicted distributions above the threshold for models A–D, respectively. doi:10.1371/journal.pone.0103899.g007

W, which separate at the boundary of the GM and NM. Habitat barriers and environmental differences, combined with geographic features, are the drivers of genetic divergences. Lineage E contains three parts: E1, E2 and E3. Thus our analyses define four ESUs that correspond to matriline W, E1, E2 and E3. Genetic and environmental data identify four historical refugia, each of which corresponds to a lineage or sublineage. Lineage W originates from a refugium in river valley near locality 27. Most populations of Lineage E originate from a refugium in the river valley near Lhasa. Two microrefugia occur in river valleys along the Himalayas. Our analyses detect a population mixture after the LGM in some localities but these occurrences do not influence our designations of ESUs. Our study highlights the importance of valleys along the Himalayas for biodiversity conservation. Based on climate models under GCC, we predict the potential distribution changes and

threats to genetic diversity of *N. parkeri*. Our pipeline, which combines DNA barcoding and SDMs, is an effective approach in conservation.

Supporting Information

Figure S1 Matrilineal genealogy of *Nanorana parkeri* based on BI analyses of COI sequence data. Bootstrap proportions $\geq 70\%$ and Bayesian posterior probabilities $\geq 95\%$ were treated as strongly supported (\blacktriangledown) and bootstrap proportions $\geq 70\%$ and Bayesian posterior probabilities $\geq 90\%$ were treated as being moderately supported. Bootstrap proportions $< 70\%$ and Bayesian posterior probabilities $< 90\%$ were treated as being unsupported (*). (TIF)

Figure S2 River systems in the southern QTP.
(TIF)

Table S1 Detailed information for specimens included in this study.
(XLSX)

Table S2 Values of F_{CT} , F_{ST} and F_{SC} based on population groups suggested by SAMOVA.
(XLSX)

Table S3 Values of π and A_r for populations from each locality.
(XLSX)

Table S4 Detailed results of AUC for random and real data.
(XLSX)

References

- Rosenzweig C, Karoly D, Vicarelli M, Neofotis P, Wu Q, et al. (2008) Attributing physical and biological impacts to anthropogenic climate change. *Nature* 453: 353–357.
- Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.
- Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Phil Trans R Soc B* 359: 183–195.
- Crandall KA, Bininda-Emonds OR, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends Ecol Evol* 15: 290–295.
- Moritz C (1994) Defining 'evolutionarily significant units' for conservation. *Trends Ecol Evol* 9: 373–375.
- Hebert PD, Cywinska A, Ball SL (2003) Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci* 270: 313–321.
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecol Lett* 8: 993–1009.
- Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, et al. (2004) Extinction risk from climate change. *Nature* 427: 145–148.
- Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, et al. (2013) Predicting species distributions for conservation decisions. *Ecol Lett* 16: 1424–1435.
- Zheng D (1996) The system of physico-geographical regions of the Qinghai-Xizang (Tibet) Plateau. *Science in China (series D)* 39: 410–417.
- Zhao J, Shi Y, Wang J (2011) Comparison between Quaternary Glaciations in China and the Marine Oxygen Isotope Stage (MIS): An Improved Schema. *Acta Geographica Sinica* 66: 867–884.
- Yao T, Liu X, Wang N (2000) The question of the range of climatic change on the Qinghai-Tibetan Plateau. *Chinese Science Bulletin* 45: 98–106.
- Beebee T (2005) Conservation genetics of amphibians. *Heredity* 95: 423–427.
- Zhou W, Yan F, Fu J, Wu S, Murphy RW, et al. (2013) River islands, refugia and genetic structuring in the endemic brown frog *Rana kukunoris* (Anura, Ranidae) of the Qinghai-Tibetan Plateau. *Mol Ecol* 22: 130–142.
- Fei L, Ye C, Jiang J, Xie F, Huang Y (2005) An Illustrated Key to Chinese Amphibians. Sichuan Publishing House of Science and Technology, Chengdu.
- Che J, Zhou W, Hu J, Yan F, Papenfuss TJ, et al. (2010) Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. *Proc Natl Acad Sci USA* 107: 13765–13770.
- Che J, Hu J, Zhou W, Murphy RW, Papenfuss TJ, et al. (2009) Phylogeny of the Asian spiny frog tribe Paini (Family Dicroglossidae) sensu Dubois. *Mol Phylogenet Evol* 50: 59–73.
- Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.
- Che J, Chen H, Yang J, Jin J, Jiang K, et al. (2012) Universal *COI* primers for DNA barcoding amphibians. *Mol Ecol Resour* 12: 247–258.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876–4882.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
- Brandley MC, Schmitz A, Reeder TW (2005) Partitioned Bayesian Analyses, Partition Choice, and the Phylogenetic Relationships of Scincid Lizards. *Syst Biol* 54: 373–390.
- Brown JM, Lemmon AR (2007) The Importance of Data Partitioning and the Utility of Bayes Factors in Bayesian Phylogenetics. *Syst Biol* 56: 643–655.
- Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24: 581–583.
- Swofford DL (2003) PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods) Version 4.0b10. Sunderland, MA: Sinauer Associates.
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57: 758–771.
- Bandelt HJ, Forster P, Rohlf A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16: 37–48.
- Polzin T, Daneshmand SV (2003) On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* 31: 12–20.
- Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, et al. (2012) Improving the Accuracy of Demographic and Molecular Clock Model Comparison While Accommodating Phylogenetic Uncertainty. *Mol Biol Evol* 29: 2157–2167.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29: 1969–1973.
- Rambaut A, Drummond A (2007) Tracer v1.4. Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Dellicour S, Mardulyn P (2013) SPADS 1.0: a toolbox to perform spatial analyses on DNA sequence data sets. *Mol Ecol Resour*: in press.
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* 11: 2571–2581.
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.
- Excoffier L, Lischer H (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- Fu Y (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.
- Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* 9: 552–569.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190: 231–259.
- Phillips SJ, Dudik M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31: 161–175.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25: 1965–1978.
- Merow C, Smith MJ, Silander JA (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36: 1058–1069.
- Smith AB (2013) On evaluating species distribution models with random background sites in place of absences when test presences disproportionately sample suitable habitat. *Divers Distrib* 19: 867–872.
- Holt RD (2003) On the evolutionary ecology of species' ranges. *Evol Ecol Res* 5: 159–178.
- Peterson A, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. *Science* 285: 1265–1267.
- Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu Rev Ecol Evol Syst* 36: 519–539.
- Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, et al. (2006) The community climate system model version 3 (CCSM3). *J Clim* 19: 2122–2143.
- Hasumi H, Emori S (2004) K-1 coupled GCM (MIROC) description. Center for Climate System Research, University of Tokyo, National Institute for

- Environmental Studies, Frontier Research Center for Global Change, Tokyo, Japan.
52. Li S, Xu L, Guo Y, Qian W, Zhang G, et al. (2007) Change of annual precipitation over Qinghai-Xizang Plateau and sub-regions in recent 34 Years. *Journal of Desert Research* 27: 307–314.
 53. Wang H, Laqiong, Sun K, Lu L, Wang Y, et al. (2010) Phylogeographic structure of *Hippophae tibetana* (Elaeagnaceae) highlights the highest microrefugia and the rapid uplift of the Qinghai-Tibetan Plateau. *Mol Ecol* 19: 2964–2979.
 54. Jia D, Liu T, Wang L, Zhou D, Liu J (2011) Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biol J Linn Soc* 102: 37–50.
 55. Wang L, Wu Z, Bystriakova N, Ansell SW, Xiang Q, et al. (2011) Phylogeography of the Sino-Himalayan Fern *Lepisorus clathratus* on the Roof of the World? *PLoS one* 6: e25896.
 56. Wang L, Abbott R, Zheng W, Chen P, Wang Y, et al. (2009) History and evolution of alpine plants endemic to the Qinghai-Tibetan Plateau: *Aconitum gymnantrum* (Ranunculaceae). *Mol Ecol* 18: 707–721.
 57. Yang F, Li Y, Ding X, Wang X (2008) Extensive population expansion of *Pedicularis longiflora* (Orobanchaceae) on the Qinghai-Tibetan Plateau and its correlation with the Quaternary climate change. *Mol Ecol* 17: 5135–5145.