Current Zoology, 2024, **70**, 150–162
https://doi.org/10.1093/cz/zoad008
Advance access publication 14 March 2023
Original Article

The Silk roads: phylogeography of Central Asian dice snakes (Serpentes: Natricidae) shaped by rivers in deserts and mountain valleys

Daniel Jablonski[a,*], Konrad Mebert[b], Rafaqat Masroor[c], Evgeniy Simonov[d], Oleg Kukushkin[e,f], Timur Abduraupov[g], and Sylvia Hofmann[h,i,*]

[a]Department of Zoology, Comenius University in Bratislava, Bratislava, Slovakia,
[b]Global Biology, Birr, Switzerland,
[c]Pakistan Museum of Natural History, Shakarparian, Islamabad, Pakistan,
[d]Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia,
[e]T. I. Vyazemski Karadag Scientific Station—Nature Reserve—Branch of A.O. Kovalevsky Institute of Biology of the Southern Seas, Theodosia, Crimea,
[f]Zoological Institute of the RAS, Saint Petersbourg, Russia,
[g]Institute of Zoology, Academy of Sciences of the Republic of Uzbekistan, Yunusabad, Tashkent, Uzbekistan,
[h]Museum Koenig Bonn, LIB—Leibniz Institute for the Analysis of Biodiversity Change, Bonn, Germany, and
[i]UFZ – Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstrasse 15, 04318 Leipzig, Germany

[*]Address correspondence to Daniel Jablonski. E-mail: daniel.jablonski@uniba.sk; Sylvia Hofmann. E-mail: s.hofmann@leibniz-lib.de.
Handling editor: Jing Che

Abstract

Influenced by rapid changes in climate and landscape features since the Miocene, widely distributed species provide suitable models to study the environmental impact on their evolution and current genetic diversity. The dice snake *Natrix tessellata*, widely distributed in the Western Palearctic is one such species. We aimed to resolve a detailed phylogeography of *N. tessellata* with a focus on the Central Asian clade with 4 and the Anatolia clade with 3 mitochondrial lineages, trace their origin, and correlate the environmental changes that affected their distribution through time. The expected time of divergence of both clades began at 3.7 Mya in the Pliocene, reaching lineage differentiation approximately 1 million years later. The genetic diversity in both clades is rich, suggesting different ancestral areas, glacial refugia, demographic changes, and colonization routes. The Caspian lineage is the most widespread lineage in Central Asia, distributed around the Caspian Sea and reaching the foothills of the Hindu Kush Mountains in Afghanistan, and Eastern European lowlands in the west. Its distribution is limited by deserts, mountains, and cold steppe environments. Similarly, Kazakhstan and Uzbekistan lineages followed the Amu Darya and the Syr Darya water systems in Central Asia, with ranges delimited by the large Kyzylkum and Karakum deserts. On the western side, there are several lineages within the Anatolia clade that converged in the central part of the peninsula with 2 being endemic to Western Asia. The distribution of both main clades was affected by expansion from their Pleistocene glacial refugia around the Caspian Sea and in the valleys of Central Asia as well as by environmental changes, mostly through aridification.

Key words: biogeography, colonization, Eurasia, genetic diversity, mitochondrial DNA, Paratethys, refugia, water snakes.

The region from Anatolia to Central Asia played an important role in the evolution of the current diversity of reptiles in the Western Palearctic (Sindaco and Jeremcenko 2008; Sindaco et al. 2013). Whereas the Anatolian area is well-studied and provides several biogeographic hypotheses on the evolution of regional biota (e.g., Bilgin 2011), Central Asia (countries east of the Caspian Sea from Kazakhstan to Pakistan in the south and north-western China and Mongolia in the east) remains comparatively less explored. Several scenarios have been presented for the diversification of herpetofauna in Central Asia (e.g., Solovyeva et al. 2018; Asadi et al. 2019; Dufresnes et al. 2019; Zhou et al. 2019), suggesting that the diversity and expansion of various species in the region were linked with the Mid-Miocene climatic transition characterized by the rapid cooling and progressing aridification. Such diversification has been driven also by fluctuations of the Paratethys Sea which responded to the Messinian Salinity Crisis (Van Baak et al. 2017), one of the major events shaping the evolution, diversity, and distribution of the current herpetofauna in the Western Palearctic (Poulakakis et al. 2015). Further population divergence was formed by the Pliocene and Pleistocene climatic oscillations.

As the Eastern Paratethys (Euxinic-Caspian basin, semi-closed or closed) was present in the area from the current Black Sea to the mountains of Central Asia (Popov et al. 2006; Krijgsman et al. 2019), it raises the question of whether it affected the evolution of reptiles requiring aquatic habitats. The widely distributed dice snake, *Natrix tessellata* (Laurenti, 1768), is a suitable model to answer such a question. This species is currently distributed from Central Europe, through

Received 2 October 2022; accepted 2 March 2023

Egypt and the Middle East, Central Asia, to western China. Previous research on the phylogeny and phylogeography of the species detected 9 mitochondrial lineages with the basal radiation occurring in Western Asia (Guicking et al. 2002; 2009; Guicking and Joger 2011; Kyriazi et al. 2013). It was assumed that the diversification within the species is linked to the major aridification and cooling that occurred during the Tortonian period of the Miocene (Kyriazi et al. 2013) and progressed with the late Miocene aridification events, primarily the Messinian Salinity Crisis. This is paralleled by the evolution of other animals in that region (Poulakakis et al. 2015), and further supported by similar results from the closely related grass snake *Natrix natrix* sensu lato (Fritz et al. 2012; Kindler et al. 2013, 2017). In contrast, the major clades of *N. tessellata* originated in the Central Palearctic (Western and Central Asia), whereas those for the *N. natrix* sensu lato have their roots in the Western Palearctic realm (northern Africa and Western and Central Europe).

In contrast to the large effort in the last decade for understanding the evolution and biogeography of the snake fauna in the Western Palearctic, the genetic diversity in Asian dice snake populations was only partially understood. Rastegar-Pouyani et al. (2017) focused on the genetic diversity and distribution of lineages found in Iran, however, without discussing biogeographic consequences. Asztalos et al. (2021) presented an updated view on the phylogeography of *N. tessellata* based on previous and new samples from Anatolia and compared the phylogeographic structure and local hybridizations with *N. natrix*. These authors also found several cases of hybridization between both studied species. Nevertheless, a detailed phylogeographic study on populations from Central Asia and related populations is missing. Thus, we aimed to fill these gaps to provide insights into 1) the phylogeographic structure of the species in Central Asia and adjacent regions; 2) the geographic origin of detected lineages, and their possible refugia; and 3) conditions that affected the colonization and distribution in the arid and mountainous areas of Central Asia.

Material and Methods

Data collection

We obtained DNA data from 33 new samples of *N. tessellata* from Western and Central Asia. Tissues were transferred into absolute ethanol and stored at –20°C. DNA was isolated using the DNAeasy Tissue Kit (QIAGEN), following the manufacturer's protocols, and stored at –20°C. The material was obtained through fieldwork in areas that were not covered in previous studies (e.g., Afghanistan, Kyrgyzstan, Pakistan, and Tajikistan). Additional data were from museum collections and compiled with published sequences related to the studied region (Guicking et al. 2006, 2009; Kyriazi et al. 2013; Rastegar-Pouyani et al. 2017; Asztalos et al. 2021; Supplementary Table S1).

Amplification and sequencing

The complete mitochondrial DNA cytochrome *b* gene (1,200 bp; *cytb*) was amplified with primers L14724NAT (Guicking et al. 2002) and H16064 (Burbrink et al. 2000; modified by De Queiroz et al. 2002) using the following PCR program: 7 min denaturing step at 94°C followed by 40 cycles of denaturing for 40 s at 94°C, primer annealing for 30 s at 46–50°C, and elongation for 1 min at 72°C, with a final 7 min elongation step at 72°C. The same primers were applied for sequencing. Cytochrome *b* is the most common marker used in molecular phylogenetic analyses of reptiles, including studies on Eurasian snakes (Guicking et al. 2009; Jablonski et al. 2019). In case of lower quality DNA of old tissue samples (e.g., sample from specimen ZFMK 95022 from Kunduz, Afghanistan) we used universal *cytb* primers CB1—CCATCCAACATCTCAGCATGA and CB2—CCCTCAGAATGATATTTGTCC (modified after Kocher et al. 1989) to amplify a fragment of approximately 350 bp length. PCR products were purified using the ExoSAP-IT enzymatic clean-up (USB Europe GmbH, Staufen, Germany; manufacturer's protocol). The sequencing was performed by Macrogen Europe Inc. (Amsterdam, The Netherlands; http://www.macrogen-europe.com), and new sequences were deposited in GenBank under accession numbers OQ122168–OQ122200 (Supplementary Table S1).

Phylogenetic analyses

DNA sequences were manually checked, aligned, and inspected using Sequencher 5.4 and BioEdit 7.0.9.0 (Hall 1999). No stop codons were detected when the sequences were translated using the vertebrate mitochondrial genetic code in the program DnaSP 6.00 (Rozas et al. 2017). The same program was used to calculate uncorrected *p*-distances among the main clades, and to estimate the number of haplotypes (h) and nucleotide diversity (π). The best-fit codon-partitioning schemes and the best-fit substitution models were selected using PartitionFinder 2 (Search algorithm: all, linked branch lengths; Lanfear et al. 2017), according to the Bayesian information criterion. Phylogenetic trees for the dataset were inferred using the Bayesian approach (BA) and maximum likelihood (ML) by MrBayes 3.2.6 (Ronquist et al. 2012) and RAxML 8.0. (Stamatakis 2014), respectively. The best-fit substitution model with each codon position treated separately for the BA analysis was as follows: HKY + I + G (first and second positions), GTR + G (third position), while it was GTR + I + G in each codon position in the ML analysis. The first 20% of trees were discarded as the burn-in after inspection for stationarity of log-likelihood scores of sampled trees in Tracer 1.7.1 (Rambaut et al. 2018; all parameters had an effective sample size [ESS] of > 200). A majority-rule consensus tree was drawn from the post-burn-in samples and posterior probabilities were calculated as the frequency of samples recovering any clade. The ML clade support was assessed by 1,000 bootstraps. Nodes with posterior probability/bootstraps values ≥0.95/≥70 were considered as moderate or well-supported.

For the dating approach, we complemented our new *cytb* sequences with data available in GenBank, including all sequences of *N. tessellata* used in Guicking et al. (2006, 2009), and selected sequences from Kyriazi et al. (2013), Rastegar-Pouyani et al. (2017), and Asztalos et al. (2021) (Supplementary Table S1 and Figure S1). The resulting data set consisted of 251 sequences (1,117 bp). The time-calibrated BA was performed based on the codon-partitioned data in BEAST 2 v.2.7.0 (Bouckaert et al. 2019) using the bModelTest (Bouckaert and Drummond 2017) package in BEAST 2 to infer the nucleotide substitution models during the Markov chain Monte Carlo (MCMC) analysis. To ensure that the tree prior does not have a negative impact on molecular dating (Ritchie et al., 2017), we conducted analyses using 2 different diversification processes, Yule (pure birth) and a birth–death.

Both models were compared using the Bayes factor (BF). The marginal likelihoods for the BF calculations were estimated based on the steppingstone (ss; Xie et al. 2011) and path sampling (ps; Lartillot and Philippe, 2006) methods using in BEAST 2 with 100 million generations, a chain length of 1 million, and 100 path steps. Statistical support was then evaluated via 2lnBF using the ps/ss results sensu Kass and Raftery (1995). The dating approach followed Kyriazi et al. (2013) using "external" calibration age constraints and setting the prior on the calibration nodes so that the youngest age of the distribution corresponded to the youngest possible age at which that lineage existed using a log-normal distribution. Briefly, 6 fossil records were implemented: C1 (earliest *Coluber* and *Masticophis* fossils; mean = 0.0, *SD* = 0.843, offset = 11.0), C2 (earliest *Salvadora* fossil; mean = 0.0, *SD* = 0.843, offset = 20.0), C3 (earliest *Lampropeltis* fossil; mean = 0.0, *SD* = 0.843, offset = 15.0), C4 (earliest *Pantherophis* fossil; mean = 0.0, *SD* = 0.843, offset = 16.0), C5 (earliest *Thamnophis* fossil; mean = 14.0, *SD* = 4.70, offset = 13.4), and C6 (the first fossil appearance of the tribe Thamnophinii; mean = 19.0, *SD* = 4.75, offset = 18.8). Eight runs were performed each with 100 million generations, and a thinning range of 10,000. Replicate runs were then combined with BEAST 2 LogCombiner v.2.7.0 by resampling logs and trees from the posterior distributions at a lower frequency and using a burn-in of 10% for each dataset, resulting in a final set of approximately 18,000 trees. Convergence and stationary levels were verified with Tracer v.1.7.1 (Rambaut et al. 2018). We annotated the tree information with TreeAnnotator v.2.6.7 and visualized it with FigTree v.1.4.2.

We also constructed TCS haplotype networks for each mtDNA phylogenetic lineage detected in the studied region. For this, we included only long-length sequences (>900 bp). We analyzed all lineages except Anatolia III which was represented by one sequence only. To build networks we used the software PopArt 1.7 (http://popart.otago.ac.nz; Leigh and Bryant 2015) with the incorporated 95% connection limit. This approach allows us to present intraspecific evolution and thus, better recognizes relationships between populations (Posada and Crandall 2001).

Demography

The past population dynamics were estimated using the Bayesian skyline plot (BSP; Drummond et al. 2005), as implemented in BEAST 2 v.2.7.0. This method computes the effective population size through time directly from sampled sequences. The BSPs were applied to the Caspian, Kazakhstan, and Uzbekistan lineages of the Central Asiatic clade and the Anatolia IV lineage of the Anatolia clade, that is, lineages that had enough sequences. We used a uniform prior for the mean substitution rate for the *cytb* with the initial value of 0.00736 mutations per site/Mya taken from the final molecular clock analysis (see above). Preliminary analyses were run using both a strict molecular clock and the uncorrelated lognormal relaxed molecular clock provided similar data that were in accordance with the published estimations. Given that the parameter of the standard deviation of the uncorrelated lognormal relaxed clock was close to zero, the final analyses were run enforcing the strict molecular clock model. The best-fitting partitioning scheme was estimated using PartitionFinder 2 (all codon positions being treated together as one partition), and the HKY substitution model was selected as the best-fitting model. The final BSP analysis was run in duplicates to check for consistency between runs, each run for 10–50 million generations and sampled every 1,000 generations. Convergence, ESS > 200, stationarity, and the appropriate number of generations to be discarded as burn-in (10%) were assessed using Tracer v.1.7.1. with the maximum time as the mean of the root height parameter.

Ancestral area estimation in continuous space

The ancestral areas of the main detected *N. tessellata* lineages and the spatial-temporal patterns of diffusion throughout their distribution range were estimated using the Bayesian phylogeographic analysis in continuous space implemented in BEAST v.1.10.4 (Lemey et al. 2010). We performed separate analyses with the same settings for each main clade identified by the previous phylogenetic analysis, to avoid any potential bias caused by population structure (Heller et al. 2013; Chiocchio et al. 2021). We applied a Yule tree prior, the "Cauchy" model for spatial diffusion, and a strict molecular clock model with 1.35% sequence divergence per million years (Guicking et al. 2006). Geographical coordinates were provided for each sequence, applying a jitter of ±0.001 to duplicated coordinates. Analyses were run for 200 million generations, sampling every 20,000 generations. The convergence of the MCMC chains was inspected using Tracer v.1.7.1 to ensure adequate mixing and convergence. Finally, the sampled trees were annotated using TreeAnnotator v.2.6.2 and the final tree was analyzed in SpreaD3 v.1.0.7 (Bielejec et al. 2016) to visualize the ancestral area for each lineage.

Occurrence data species distribution modeling

We performed species distribution modeling (SDM) to illustrate the past and recent geographical distribution of *N. tessellata*. Grids of 19 standard bioclimatic variables were downloaded from the WorldClim (http://www.worldclim.org) and CHELSA (https://chelsa-climate.org/) databases for the Last Glacial Maximum (LGM) (~22 Kya; Karger et al. 2023), Mid-Holocene (~6 Kya; Fordham et al. 2017; Brown et al. 2018), and current climate. All layers were clipped to the extension of the known *N. tessellata* range and projected to WGS84 using ArcGIS 10.8 (ESRI, Redlands, CA, USA). To reduce the autocorrelation of climate data we removed highly correlated variables based on the current climate data sets and Pearson's correlation coefficients ($r > 0.7$) using the python script SDMtoolbox v.2.5 (Brown 2014) available for ArcGIS. The final bioclimatic data sets comprised 8 variables: BIO1 = annual mean temperature, BIO2 = mean diurnal temperature range, BIO4 = temperature seasonality, BIO8 = mean temperature of wettest quarter, BIO12 = annual precipitation, altitude (CHELSA data), BIO14 (WorldClim data) = precipitation of driest month, BIO15 = precipitation seasonality, and BIO19 = precipitation of coldest quarter.

Based on the recent climate data set, for these variables, we carried out a principal components analysis (Figure 5D) with SDMtoolbox and used the Eigenvalues of the resulting 3 components (88.79, 7.42, and 3.80) to assess the weighted climatic heterogeneity (Figure 5E) across the area of interest. To eliminate spatial clusters of localities we spatially filtered our presence data by Euclidian distances (min 2 km, max 25 km; 5 distance classes) according to climate heterogeneity using the rarefying module in SDMtoolbox; as a result, 3 points were excluded. Prior to the SDMs, background points were selected by a buffered minimum-convex polygon based on known occurrences using a buffer distance of 200 km. The

final 2 models were generated with MaxEnt v.3.4.3 (Phillips et al. 2004, 2006) which is adequate for analyzing presence-only data and based on the principle of maximum entropy. Model performance and the importance of the environmental variables to the model were assessed using the mean area under the curve (AUC) of the receiver operating characteristics (Hanley and McNeil 1982; Robin et al. 2011), and jack-knife testing; models with AUC values above 0.75 are considered potentially informative (Elith et al. 2006), good between 0.8 and 0.9, and excellent for AUC between 0.9 and 1 (Préau et al. 2018). For projecting the MaxEnt models based on past climates, we used again the SDMtoolbox in ArcGIS.

Results

Molecular phylogeny, time divergence, and phylogeography

The ML and BA gene trees based on a data matrix of 194 DNA sequences (190 without outgroup taxa) yielded almost identical tree topologies with 7 major, well-supported clades of *N. tessellata* (except the Greece clade in ML and Jordan clade in BA tree) (Figure 1; for BA trees see Supplementary Figures S1 and S2). These clades are named as follows: Iran, Jordan, Greece, Crete, Europe, Anatolia (=Turkey lineage), and Central Asia (modified from Guicking et al. 2009). The last two, sister clades comprise 4 lineages (I–IV) in the Anatolia clade (Anatolia III is represented by a single sequence from Armenia) and 3 lineages in Central Asia clade (Caspian [= Caucasus modified from Guicking et al. 2009], Kazakhstan, and Uzbekistan; Figure 1). In the Anatolia clade, lineages Anatolia I and Anatolia II (both endemic to Anatolia), constitute the sister position, whereas the widely distributed lineage Anatolia IV shows a sister relationship with the Armenian lineage (Anatolia III). In the Central Asia clade, the Kazakhstan lineage is sister to both Caspian and Uzbekistan lineages that diverged later than the first one (see Figure 1A, B and Supplementary Figures S1 and S2).

We performed molecular dating under Yule and a birth–death model. Both approaches provided different but comparable results (Supplementary Table S2). For the final phylogenetic inference, we applied a birth–death process because this prior has been shown to be more suitable for shallow and balanced phylogenies (Brown and Yang 2010), while the Yule tree prior is more suitable for between species branching (Drummond et al. 2007). Moreover, mean dating results under birth–death model are more comparable with other published data (Kyriazi et al. 2013). Overall, the age estimations under the birth–death process are younger than those inferred under the Yule model. The birth–death model also strongly outperformed the Yule model (2lnBF > 20); Table S3). Hence, we preferred the results under the birth–death process to the Yule model. Molecular dating estimations in *N. tessellata* suggest divergence from ancestral *N. tessellata* begun in the late Miocene with the expected split of the Iran clade at 9.6 Mya (11.6–7.1 Mya of 95% highest posterior density—HPD), followed by the Greece clade at 7.8 Mya (9.6–5.9), Jordan clade at 7.1 Mya (8.3–5.4), and finally Crete plus Europe clades at 6.0 Mya (7.5–4.5). The latter 2 clades diverged from each other probably between the end of the Pliocene and the middle of the Pleistocene approximately 2.3 Mya (3.1–1.3 Mya), whereas Anatolia and Central Asia clades diverged earlier around 3.7 Mya (4.5–2.6 Mya). In Anatolia and Central Asia, 4 and 3 lineages were detected, respectively, that is, Anatolia I–IV and Caspian, Kazakhstan, and Uzbekistan. These lineages started their divergence during the early Pleistocene with the initial split of Anatolia lineages around 2.4 Mya (3.0–1.6 Mya) and around 2.5 Mya (3.1–1.6 Mya) among Central Asia lineages (Figure 1A, B and Supplementary Figure S2). The phylogenetic positions of clades have been well resolved and supported, except for the weaker support and uncertain placement of the Jordan clade (Figure 1A).

The genetic diversity of Anatolia and Central Asia clades shows a clear geographical pattern (Figure 2A) with well-recognizable internal structures (Figure 1B and Supplementary Figure S2). In the western part of the range (Anatolia, Transcaucasia, and Black Sea region) 5 lineages are present (Anatolia I–IV, Caspian), and 4 lineages are in the region south and east of the Caspian Sea (Anatolia IV, Caspian, Kazakhstan, and Uzbekistan). Lineages Anatolia I and II are geographically restricted with a distribution in central-south (Cilicia) and western Türkiye where they approach the Europe clade of *N. tessellata*. The distinct lineage from Armenia (sequence ID 2986) originates from an area where geographically widespread lineages Anatolia IV and Caspian meet. These 2 lineages also meet in northern Iran, that is, along the Talysh and Alborz Mountains where they possibly form a contact zone with the deeply diverged Iran clade (Figure 2A). The geographically most widespread Caspian lineage is also present as a distant single record in the southern Zagros Mountains (see Discussion). Similarly, the Anatolian IV lineage that continuously represents records from Cyprus to Central and Eastern Anatolia, and across Transcaucasia, has also been detected as an isolated record in north-central Iran (Figure 2A). The Caspian lineage occupies the area around the Caspian Sea and north of it, approximately 3,000 km from central Kazakhstan and Afghanistan to Eastern Europe (Crimean Peninsula and Southern Bug River, Ukraine). Southeast of the Caspian Sea, this lineage continues east along the Kopet Dag Mountains limited north by the inhospitable Karakum Desert in Turkmenistan, and equally south by the arid areas of Dasht-e-Kavir in Iran and Sistan Basin in southern Afghanistan (Figure 2A). The Caspian lineage reaches its easternmost site in central Afghanistan, probably following the Hari River. Finally, the Caspian lineage possibly meets the Kazakhstan and Uzbekistan lineages north to the Aral Lake. The Kazakhstan lineage is restricted to the Syr Darya basin (northern Uzbekistan, southern Kazakhstan, and northern and central Kyrgyzstan), and geographically limited due to the Kyzylkum Desert and mountains of Tajikistan. Adjacent to the south, the Uzbekistan lineage is linked to the Amu Darya basin (western and southern Uzbekistan, Tajikistan, and northern Afghanistan) and geographically constrained by the Kyzylkum Desert in the north and the Karakum Desert in the south. In Afghanistan, the Uzbekistan lineage possibly meets the Caspian lineage somewhere around Paropamisus and foothills of the Hindu Kush Mountains. The Uzbekistan lineage also expanded east along Amu Darya and Panj river systems into the Pakistani Hindu Kush close to Karakoram (Figure 2A) and the Indus River (Gilgit area; northern Pakistan), which constitutes the border between the Palearctic and the Oriental zoogeographical realms. Noteworthy, the Uzbekistan lineage is characterized by a high number of missing/extinct haplotypes and the highest nucleotide variability (π = 1.2%) compared to the

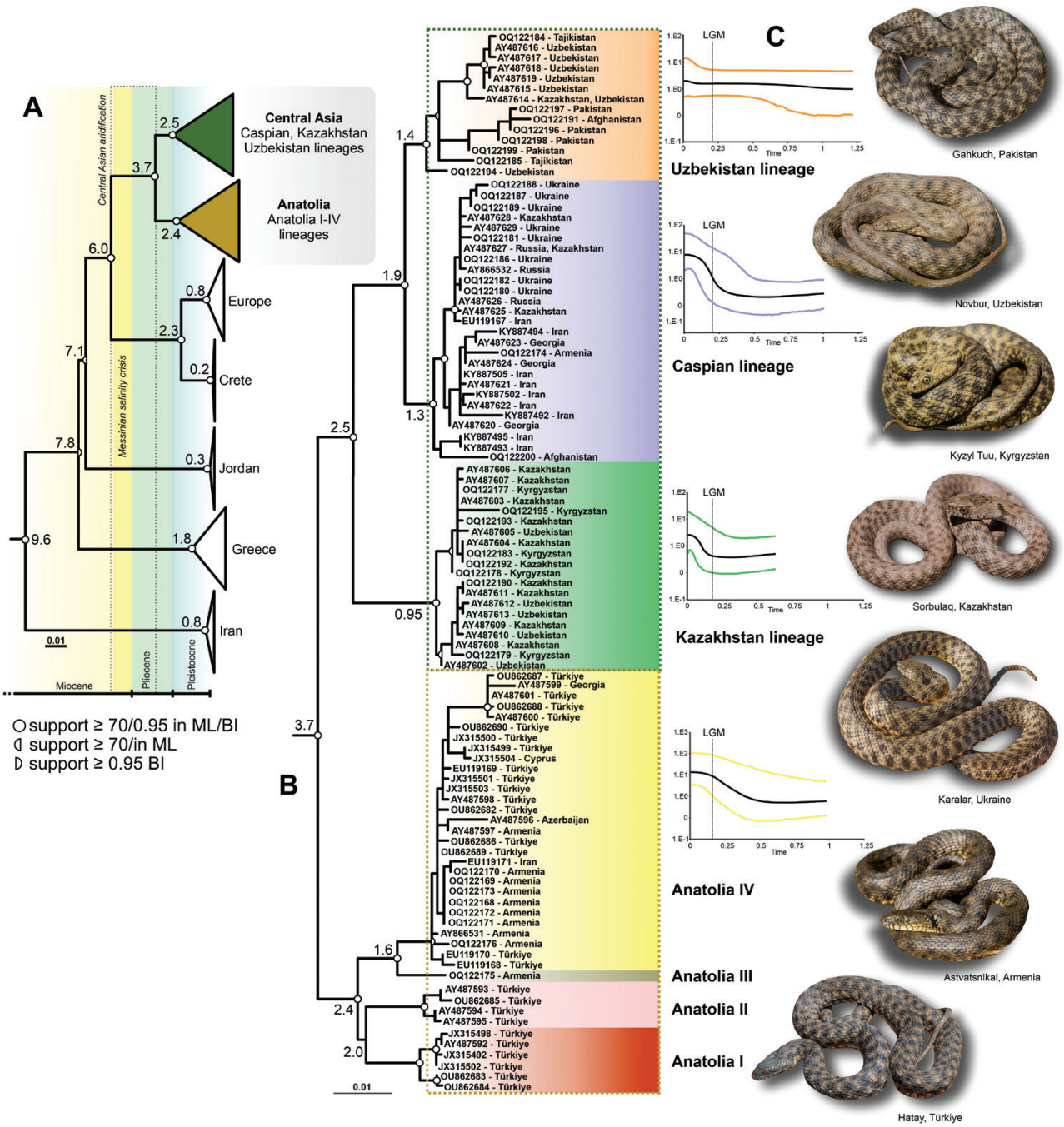

Figure 1 The maximum likelihood tree reconstruction of *Natrix tessellata* (1,117 bp; Supplementary Table S1) with time divergences based on the molecular clock analysis (A), and the inset to relationships between Anatolia and Central Asia clades (B). Numbers at nodes represent the expected time of the divergence. Terminal branch labels consist of the sample ID (new material) or GenBank accession number, and the name of the country or region of origin (see also Figure 1 and Supplementary Table S1). (C) Bayesian skyline plots showing the historical demography for the investigated lineages representing enough number of sequence data: the central line shows the mean value of the population size ($N_e \times \tau \times \mu$; where N_e is the effective population size, τ is the generation length in units of time [substitutions/site], and μ is the mutation rate) on the logarithmic scale. LGM line indicates the time of the Last Glacial Maximum. Inset photographs: Daniel Jablonski and M. M. Beskaravaynyi.

other lineages, which are less structured with a nucleotide variability equal to or lower than 0.5% (Supplementary Figure S2).

The average of genetic distances among studied lineages of the *N. tessellata* varied from 2.1% (Kazakhstan vs. Uzbekistan lineages) to 5.0% (Anatolia II vs. Caspian lineages). Intralineage divergence reached the highest value of 0.9% in the Anatolia I lineage (Table 1). The average of genetic distances between Anatolia and Central Asia clades was 3.9%.

Demographic history

The BSP analysis (Figure 1C) showed different population demography of the investigated lineages. The mean value of the population growth (N_e) was detected in all lineages except

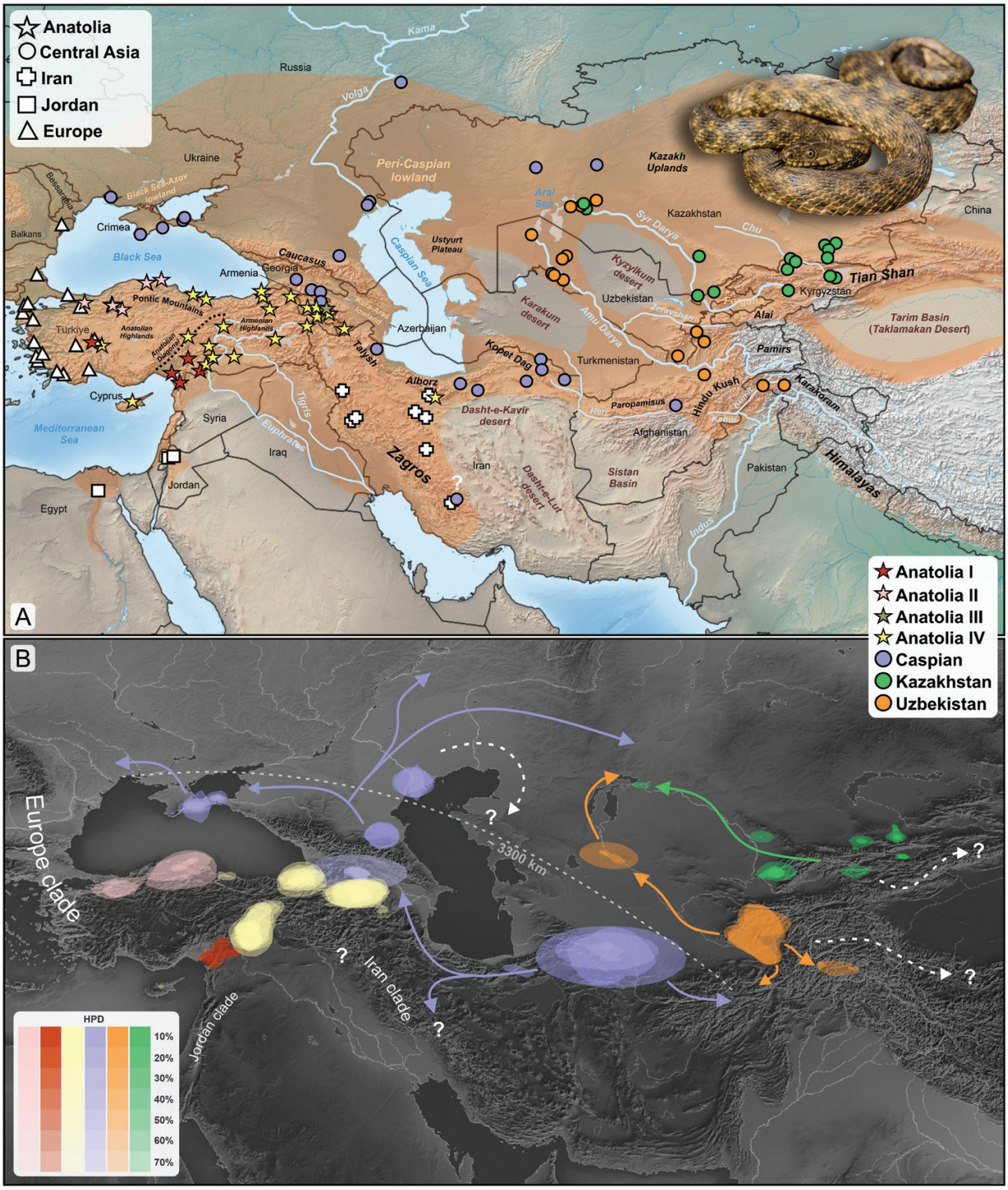

Figure 2 Upper panel (A): Geographic origin of sequences of the Anatolia and Central Asia clades of *Natrix tessellata* used in our study. Colors correspond to the main phylogenetic lineages recovered in our analysis (Figure 1; for locality details see Supplementary Table S1). Neighboring clades (Iran, Jordan, Europe, sensu Guicking et al. 2009) are indicated by different symbols. The distribution range of the species is highlighted in light orange. The question mark denotes uncertain origin of sequence KY887502. The pictured individual originates from Kazarman, Kyrgyzstan (Kazakhstan lineage). Lower panel (B): Ancestral areas of the genetic lineages of *N. tessellata* from Anatolia and Central Asia. Polygons represent regions with 10–70% highest posterior density (HPD) of the ancestral areas. The hypothetical colonization routes Caspian, Kazakhstan, and Uzbekistan lineage indicating by color arrows. Question marks indicate unknown genetic affiliation in unstudied areas and white arrows hypothetical origins. The map was drawn using QGIS 3.20. (https://qgis.org). Inset photograph: Daniel Jablonski.

Table 1 Average uncorrected *p*-distances (%) calculated among the cytochrome *b* sequences of the main lineages of Anatolia and Central Asian clades of *Natrix tessellata*. In diagonal (italics) are the average intraclade *p*-distances. The highest value of *p*-distance between Anatolia 2 and Caspian lineage is highlighted in bold

P-distance (%)	Anatolia I ($n = 6$)	Anatolia II ($n = 4$)	Anatolia III ($n = 1$)	Anatolia IV ($n = 29$)	Kazakhstan ($n = 20$)	Caspian ($n = 27$)	Uzbekistan ($n = 14$)
Anatolia I	*0.9*						
Anatolia II	2.3	*0.4*					
Anatolia III	2.7	2.8	0				
Anatolia IV	2.6	2.9	2.6	*0.4*			
Kazakhstan	3.9	3.9	2.8	4.3	*0.4*		
Caspian	3.9	**5.0**	4.1	4.0	3.2	*0.7*	
Uzbekistan	3.7	3.6	2.6	4.1	2.1	4.0	*0.7*

Uzbekistan where certain stability was observed before and after the LGM. The sign of the population growth started before the LGM is visible in the Anatolia IV lineage (mean 40 Kya) and the Caspian lineage (35 Kya), and after the LGM in the Kazakhstan lineage (19 Kya).

Ancestral areas, refugia, and SDM

The ancestral areas of the lineages (all except Anatolia III) were restricted to distant regions across and east of Anatolia, the Black and Caspian Seas (Figure 2B), and river valleys or lower elevated areas of Central Asia. For the widespread Caspian lineage, the ancestral areas were detected in the Kopet Dagh area of north-eastern Iran and southern Turkmenistan (time to the most recent common ancestor—TMRCA—95% HPD: 1.2–0.2 Mya), in the Transcaucasian region (TMRCA: 0.9–0.4 Mya), and in the area north of Black Sea and Caucasus (Crimea; TMRCA: 0.69–0.3 Mya). Further to the east, putative ancestral areas of the Uzbekistan lineage were mainly located along the Amu Darya, Pamirs, and the Hindu Kush (TMRCA: 1.7–0.82 Mya), while the Kazakhstan lineage had its ancestral areas probably in river valleys of the Tian Shan in southern Kazakhstan and northern Kyrgyzstan (TMRCA: 1.3–0.4 Mya; Figure 2B and Supplementary Figure S2).

The ancestral area estimations were partially consistent with the environmental niche model projection (SDM) on climatic conditions during the LGM, representing putative glacial refugia (Figure 3A). Accordingly, the main refugia during the LGM were apparently located in the Aral-Caspian depression (Caspian lineage), especially in the Kura-Aras lowland that delimitated the Greater and Lesser Caucasus and the lowland north of the Caspian Sea, that is, Peri-Caspian lowland and in the lowlands of Syr Darya and Chu rivers (Kazakhstan lineage) in southern Kazakhstan, northern Kyrgyzstan, and in the Fergana Valley (Uzbekistan lineage). One of the refugia of the Caspian lineage was probably located in the northwestern part of the Black Sea area. The projection during the Mid-Holocene (6 Kya; Figure 3B) shows the increased areal size to the north and west along climatically suitable areas, and the present model corresponds well to the current species range (Figures 2A and 3C).

The current climate across the modeled area is spatially heterogeneous, particularly in Anatolia, the Caucasus, south of the Caspian Sea, and further east in the mountain valleys of Central Asia (Figure 3D). All these areas share similar climatic conditions (indicated by similar color, Figure 3E), and are known distributions of *N. tessellata*.

Discussion

The evolution and phylogeography of the dice snake *N. tessellata* have been previously addressed with studies by Guicking et al. (2006, 2009), Guicking and Joger (2011), Kyriazi et al. (2013), Rastegar-Pouyani et al. (2017), and Asztalos et al. (2021). We complemented them with additional samples from the species' eastern distribution, the most poorly studied regions of Central Asia. Our study also comprised the first genetic data from the barely accessible river valleys of the Paropamisus and Hindu Kush Mountains in Afghanistan and Pakistan. The results confirmed the published molecular phylogeny of *N. tessellata* (Guicking et al. 2009; Kyriazi et al. 2013) and identified at least 7 lineages within the Central Asia and Anatolia clades. Our results further support the proposed scenario of the species' origin in southwestern Asia (Guicking et al. 2009; Guicking and Joger 2011), followed by the diversification of ancestral lineages from Iran (9.6 Mya), westward to Egypt and Levant, Greece, and the rest of Europe during the Miocene, and Anatolia and Central Asia during the Pliocene. We also support the results of Kyriazi et al. (2013) that dated the basal divergence of *N. tessellata* (split of the Iranian clade) to an older age than previously reported by Guicking et al. (2009) and parallels the basal diversification within the sister species complex, *N. natrix*, around the same period (Fritz et al. 2012). This age is also consistent with early divergence times of other reptile groups with a similar distribution range in Western and Central Asia (e.g., *Phrynocephalus*; Solovyeva et al. 2018), and suggests a parallelism in the evolution of regional biota, driven by environmental changes (Central Asian aridification) since the Miocene (Guo et al. 2004). However, for a better resolution of *N. tessellata* divergence and phylogeography in space and time, it will be necessary to study multilocus dataset or genomic data.

We have evaluated and refined several contact zones between clades. The current distribution ranges of the large Anatolia and Central Asia clades meet in the Transcaucasian region, particularly in the Armenian Highlands, Lesser Caucasus Mts., and Alborz Mts. in Iran. This confirms the relevance of that area for phylogeographic studies (Tuniyev 1995; Ahmadzadeh et al. 2013; Zinenko et al. 2015; Jablonski et al. 2019, 2021; Stratakis et al. 2022). Possible contact zones between other clades of *N. tessellata* were detected in western Türkiye (Anatolia with the Europe clade), the Levant (Anatolia and Jordan clades), Iran (Anatolia, Central Asia, and Iran clades), and in the easternmost Balkans and Bessarabia in Moldova (Central Asia and Europe clades; Figure 2). These multiple contact zones in Western Asia are

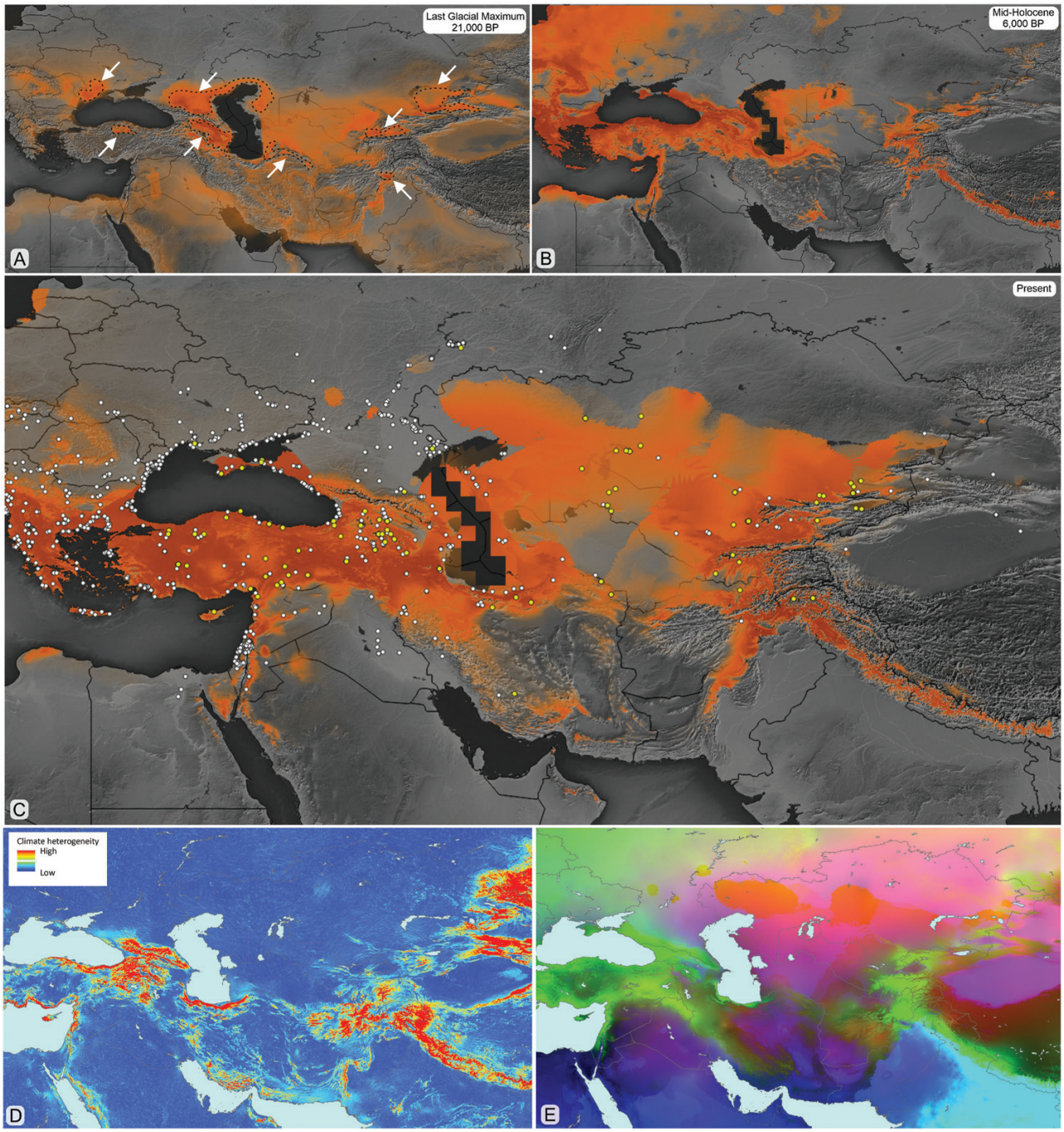

Figure 3 The environmental niche model projection and the climate heterogeneity raster based on WorldClim data. (A) Climatic conditions at the Last Glacial Maximum (LGM). The white arrows suggest examples of potential glacial refugia of the species during the LGM. (B) The model for the Mid-Holocene. (C) Present model with the layer of data distribution points (white circles) downloaded from GBIF (2022). Yellow circles represent data used for the SDM projection. (D) Warm colors depict high areas of climatic heterogeneity in the present time. (E) Principal components analysis (PCA) of filtered WorldClim variables, showing climate space: the more similar the colors the more similar values. The maps were designed in QGIS 3.20 using Min–Max as the stretching histogram.

reflected by the highest genetic variation within *N. tessellata*, because of clades having evolved separately over a long time (Figure 1A) and thus, offer a particularly suitable window for studies on gene flow, potential hybridization, and the evolutionary dynamics between clades and lineages (see Asztalos et al. 2021).

The separation of the Anatolia and Central Asia clades began approximately 3.7 Mya (Figure 1) and continued during the end of the Pliocene and beginning of the Pleistocene probably within separated refugia in Anatolia, Caucasus, and Black Sea regions, as well as lowlands and mountain valleys of Central Asia; a phylogeographic evolution similar to the partially sympatric *N. natrix* (Asztalos et al. 2021). We detected at least 4 lineages within the Anatolia clade, 2 of them (I and II) endemic to Western Asia, that is, the Anatolian region (cf. Guicking et al. 2009; Asztalos et al. 2021; and Figure 2 this study). The ancestral diversification of these lineages occurred at the beginning of the Pleistocene (2.4 Mya; Figures

1B and 2B), separately in south/western Anatolia (I and II) and Armenian highlands (III and IV; Figure 2B). This separation appears to be related to the heterogeneous topography in large parts of Anatolia, such as the biogeographic break termed Anatolian Diagonal or its derivatives (see related examples of snake biogeography; Jablonski et al. 2019; Šmíd et al. 2021), that creates a multitude of ecosystems from cold streams in tall mountains down to large humid-warm wetlands adjacent to the Mediterranean and Black Seas. Currently, dice snakes reach maximum elevations between 2,500 and 3,000 m above sea level (e.g., Yakovleva 1964; Tuniyev et al. 2011). Anatolian lineages could survive glacial periods along rocky shores of the southern Black Sea coast, where the temperature profile was more suitable compared to the adjacent highlands as demonstrated by mean annual sea surface temperatures that fluctuated between 6 and 9°C during glaciation (Wegwerth et al. 2015). These conditions resemble those that dice snakes experience today in northern relict populations (e.g., Kotenko et al. 2011, Litvinov et al. 2011), or colder regions of the Pontic Mountains, Armenian Highlands, and along the Kura-Aras river system (Figure 3A). Anatolian lineages I and IV probably survived in refugia west and east of the Euphrates River, with the lineage IV expanding during post-glacial warming westward across the Euphrates and forming a mixed distribution pattern with the southern Anatolian lineage I. This is, however, preliminary as we miss the genetic data from Iraq and Syria. Similarly, the endemic lineage Anatolia III consists of a single record from Armenia, which might represent a local population divergence restricted to Lake Sevan in Armenia, and where they now occur in sympatry with the Caspian and Anatolia IV lineages. Again, more genetic data are needed from that region, especially from Armenia, Azerbaijan, Georgia, and northern Iran, to better understand the phylogeographic patterns of *N. tessellata*. We assume that dice snakes of central to eastern Anatolia may have persisted during the coldest glacial period in open microrefugia of xerophytic steppes, semi-deserts, and south-exposed mountain slopes with plenty of solar radiation reaching the ground level, which is relevant for proper thermoregulation (Adams and Faure 1997; Mebert et al. 2013; Pickarski et al. 2015), thus experiencing further population growth and expansion (see Anatolia IV lineage, Figure 1C). A similar evolutionary diversity has been suggested in other reptile species of the Anatolian-Transcaucasian region, including snakes (Fritz et al. 2009; Sindaco et al. 2013; Hofmann et al. 2018; Jablonski et al. 2019, 2021; Asztalos et al. 2021; Šmíd et al. 2021), thus, reflecting the effects of natural barriers across those complex landscapes (Bilgin 2011). The occurrence of the Anatolia IV lineage in Cyprus suggests transmarine dispersal, for example, through floating objects, swimming (*N. tessellata* is tolerant to saline water, see Gruschwitz et al. 1999), or human-induced colonization by shipping activities in the past (Göçmen and Mebert 2011; Kyriazi et al. 2013; Burton 2021).

In Central Asia, clade diversification of *N. tessellata* began between the late Pliocene and the early Pleistocene around 2.5 Mya (Figures 1 and 2), with subsequent splits at 1.9 Mya (Caspian from Uzbekistan). The intralineage variation is comparable across lineages, suggesting similar environmental conditions that drove their mutual evolution and rapid dispersion across large corridors. This is consistent with the phylogeography of some amphibians and reptiles in Central Asia, such as the green toads (*Bufotes* spp.; Dufresnes et al. 2019), racerunners (*Eremias* spp., Guo et al. 2011) or toad-headed agamas (*Phrynocephalus* spp., Melville et al. 2009; Solovyeva et al. 2018) that adjusted to aridification, alterations in the Parathetys Basin, formation of large rivers (particularly the Amu Darya and Syr Darya), and tectonic uplifts during the Pliocene and the Pleistocene. Even though the dates listed above are in good agreement with results on time divergence within the Central Asian lineages of *N. tessellata*, the historical reconstruction of this species is complicated by the high dynamics of paleo-environments with cycles of flooding events and drainages, leading to large geophysical changes of lowland rivers, tributaries, lakes, and basins during this epoch that lasted up to 1.8 Mya. However, we can expect that the evolution of 3 major lineages in Central Asia is connected to a variety of water habitats (Caspian lineage) and the development of river systems (Kazakhstan and Uzbekistan linages), particularly the formation of deeply incised valleys in the Parathetys Basin (e.g., palaeo-Amu Darya) during the Plio-Pleistocene (Popov et al. 2006). Furthermore, vegetational changes impacted the regionally available habitats, such as the depletion of the floristic composition of vegetation, and the replacement of savannahs, forests, and forest-steppes by open steppes and deserts (Naidina and Richards 2020; Lazarev et al. 2021).

The intralineage diversity of the Caspian, Kazakhstan, and Uzbekistan lineages suggests surviving in different Pleistocene microrefugia (Figure 2B), particularly in river valleys (Fergana, Chu, Amu Darya, and Syr Darya), and along the southern slopes of Alai and Tian Shan Mountains (Kazakhstan lineage), as well as western Alai, Pamir, and Hindu Kush Mountains (Uzbekistan lineage; see Figure 3A). Whereas the Kazakhstan lineage is divided into 2 sublineages of geographically mixed populations today (Figure 1B and Supplementary Figure S3), the pattern in the Caspian lineage shows a clear structure of 3 areas that likely represent Pleistocene refugia as more land masses and their wetlands along the coasts were exposed due to sea level decrease by 20–120 m globally (e.g., Ganopolski et al. 2010). These refugia could be allocated in 1) north-eastern Iran and central Afghanistan, 2) western Iran and Transcaucasia, and 3) steppe areas north of the Black Sea (e.g., shallow brackish wetlands of the Odesa Bay and the Sea of Azov), east to lowlands north of the Caucasus and the Peri-Caspian region (Figures 2B and 3A). The Uzbekistan lineage lacks readable phylogeographic structure, even though our data indicate microrefugia in the Amu Darya Basin or even in river systems of the southern Hindu Kush that were sources of further colonization events. Comparable refugia also apply to other biotas, for example, the Central Asian green toads of the *Bufo viridis* complex (Zhang et al. 2008; Dufresnes et al. 2019), and even walnut trees (Aradhya et al. 2017).

We assume that the primary expansion of the ancestors' population of Central Asian *N. tessellata* correlated with the increased pluvial conditions in the Late Pliocene and Early Pleistocene, caused by the great Akchagylian transgression, respectively, extension, of the Caspian Basin with a reduced salinity from 5‰ to 9‰ at the beginning and 18–25‰ at the end of that epoch, and lasting from 3.6 to 1.8 Mya or 2.95–2.13 Mya, depending on the source (see Gerasimov 1976; Esin et al. 2019; Lazarev et al. 2021). In addition, the colonization of the area between the Black and the Caspian Seas was enabled through the Manych-Kerch Strait, a spillway connecting these seas in the Late Pleistocene (Svitoch 2013).

The Pleistocene and the Holocene colonization routes of *N. tessellata* in Central Asia correspond with rivers (and their drainage systems) and lakes suitable for this semi-aquatic snake (see Mebert 2011; Mebert and Masroor 2013; Mebert et al. 2013). The best example is Central Asian populations where particular lineages correspond to a specific river basin. Similar linkage has been shown in the semiaquatic colubrid snake *Thermophis baileyi*, inhabiting hot springs in the high-elevated Tibetan Plateau, where the phylogeographic pattern corresponded clearly to the drainage system (Hofmann et al. 2014).

The Caspian lineage (= Caucasus sensu Guicking et al. 2009) inhabits huge areas (ca. 3,300 km from west to east) around the Caspian Sea and north of the Black Sea, that is, the area of the former Paratethys. This shows the unprecedented colonization properties of *N. tessellata*, as they have been observed to quickly find and populate preferred habitats and build large (articles in Mebert 2011; Pauwels et al. 2020). The Caspian region with its wide shallow water areas, a large number of ground-dwelling gobiid fish (Gruschwitz et al. 1999; Tuniyev et al. 2011), and massive shoreline habitat have probably become a center of dispersion for *N. tessellata*. From there dice snakes expanded into different directions, particularly from the southeast (today eastern Iran, central or western Afghanistan) to the northwest and possibly back to Central Asia from areas north of the Caspian Sea (Figures 1A and 2B). Around the Aral Sea, it approached the Kazakhstan and Uzbekistan lineages where the exact position of their potential contact zones is currently unknown. The expansion was probably rapid as suggested by the structure of the haplotype network, exhibiting a high number of closely related haplotypes, and the BSP analysis (Figures 1 and 2 and Supplementary Figure S3). The Caspian lineage was able to reach the European continent in central-southern Ukraine and was detected even in one isolated and distant sample in southern Iran (SUHC 1842: Rastegar-Pouyani et al. 2017). However, this single record from southern Iran should be taken with caution as the material was provided by an illegal snake catcher with an unprecise, perhaps incorrect, origin from somewhere in the Fars Province (Rastegar-Pouyani pers. comm.). Hence, further investigations are required, for example, based on material from Fars Province that was studied only morphologically (Rajabizadeh et al. 2011).

The Uzbekistan and Kazakhstan lineages apparently expanded westward along the Amu Darya and the Syr Darya between Kyzylkum and Karakum deserts. The Uzbekistan lineage also crossed the Hindu Kush Mountains following rivers in north-eastern Afghanistan and reaching the Karakoram Mountains in Pakistan (Mebert and Masroor 2013; Mebert et al. 2013; and Figure 2B). The lack of signals indicating population expansion after LGM in the Uzbekistan and Kazakhstan lineages may suggest their long-term refugia (and possible bottleneck) in Central Asian Mountain ranges during Pleistocene glaciations. Given the strictly semiaquatic biology of *N. tessellata* (Mebert 2011) we suggest that the species expanded its range repeatedly during warm periods northeast along the base of the Tien Shan Mountains, around the lakes Issyuk Kul, Balkhash, and Alakol and their drainage systems. Proceeding from these regions, they could colonize the Chinese Junggar Basin, Turfan Depression, and the Tarim Basin, as records of dice snakes in north-western China indicate (Liu et al. 2011). However, these populations have not been genetically studied so far and the exact colonization route to the easternmost places of the species range is still hypothetical (Figure 2B). A relatively large number of genetic differences between haplotypes inside the Uzbekistan lineage might be triggered by the rugged and highly elevated topography of the Hindu Kush and Karakoram Mountains. It likely promoted frequent temporary isolations of populations during climatic fluctuation in the Pleistocene, which is corroborated by the lack of population growth after the LGM (Figure 1C). Especially at the edge of the species distribution in Pakistan, the haplotypes from the south-eastern-most population at Ghakuch are very distant from other populations (Supplementary Figure S3), suggesting prolonged isolation. While the exact limit of *N. tessellata* has not been established, the Indus River approximately 70 km downstream from Ghakuch, beginning east of Gilgit-Jalal Abad turns into a fast-flowing sediment-rich river, which likely is an inferior fishing ground. Another 100 km farther downstream Indus River, the presence of the ecologically very similar Asiatic water snake *Fowlea piscator* possibly prohibited the further expansion of *N. tessellata* (Mebert and Masroor 2013; Mebert et al. 2013). An assessment of the status and biogeographic history of these lineages requires a wider sampling in the area.

There is also evidence of a significantly wider distribution of the dice snake during the Early and Middle Holocene, respectively, the Holocene Climatic Optimum (HCO or Atlantic epoch). During the HCO, the northern hemisphere has been characterized by regionally higher temperatures and wetter climate (9.3–5.7 Kya; Yakovleva and Bakiev 2010; Marosi et al. 2012; Mebert et al. 2013). In such warmer periods, *N. tessellata* expanded to the north of western Eurasia (Figure 3B) outside of its contemporary distribution (Ratnikov 2009; Ratnikov and Mebert 2011). Simultaneously, the same interglacial event probably reduced the species range in Central Asia where significant desertification has developed. On the other hand, the huge Caspian Sea likely acted as a large repository for dice snakes in Central Asia, promoting expansion into adjacent tributaries, whenever conditions became suitable. The rapid colonization within a few years of various new artificial islands in the Caspian Sea, located at distances of 8–50 km from the mainland (Pauwels et al. 2020), is a perfect example of the high expansion potential of *N. tessellata*, a fact that is evident from its wide distribution range today. It highlights the high versatility of *N. tessellata*, a species that will provide us with many more fascinating attributes to study.

Acknowledgments

We thank many of our colleagues, friends, and local people for their support, material, information, or help in the field or in the laboratory. Special thanks are given to anonymous reviewers for their beneficial comments and suggestions for revised versions of the manuscript. DJ was supported by the Slovak Research and Development Agency under the contract APVV-19-0076 and by the grant VEGA 1/0242/21 of the Scientific Grant Agency of the Slovak Republic. SH was supported by the German Research Foundation (DFG, grant no. HO 3792/8-1). The work of OK was carried out within the framework of research topics of the state assignments nos. 121032300023-7 and 122031100282-2. The research of DJ in Afghanistan has been approved by the National Environmental Protection Agency of the Islamic Emirate of Afghanistan (permits for access to genetic resources nos. 12429 and 12455).

Supplementary Material

Supplementary material can be found at https://academic.oup.com/cz.

References

Adams JM, Faure H, 1997. Preliminary vegetation maps of the world since the Last Glacial Maximum: An aid to archaeological understanding. *J Archaeol Sci* **24**(7): 623–647.

Ahmadzadeh F, Flecks M, Rödder D, Böhme W, Ilgaz C et al., 2013. Multiple dispersal out of Anatolia: Biogeography and evolution of oriental green lizards. *Biol J Linn Soc* **110**: 398–408.

Aradhya M, Velasco D, Ibrahimov Z, Toktoraliev B, Maghradze D et al., 2017. Genetic and ecological insights into glacial refugia of walnut (*Juglans regia* L.). *PLoS One* **12**: e0185974.

Asadi A, Montgerald C, Nazarizadeh M, Moghaddasi A, Fatemizadeh F et al., 2019. Evolutionary history and postglacial colonization of an Asian pit viper *Gloydius halys caucasicus* into Transcaucasia revealed by phylogenetic and phylogeographic analyses. *Sci Reports* **9**: 1224.

Asztalos M, Ayaz D, Bayrakci Y, Afsar M, Tok CV et al., 2021. It takes two to tango: Phylogeography, taxonomy and hybridization in grass snakes and dice snakes (Serpentes: *Natricidae*: *Natrix natrix*, *N. tessellata*). *Vertebr Zool* **71**: 813–834.

Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A et al., 2016. SpreaD3: Interactive visualisation of spatiotemporal history and trait evolutionary processes. *Mol Biol Evol* **33**: 2167–2169.

Bilgin R, 2011. Back to the suture: The distribution of intraspecific genetic diversity in and around Anatolia. *International J Mol Sci* **12**: 4080–4103.

Bouckaert RR, Drummond AJ, 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evol Biol* **17**: 1–11.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M et al., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comp Biol* **15**: e1006650.

Brown JL, 2014. SDMtoolbox: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods Ecol Evol* **5**: 694–700.

Brown JL, Hill DJ, Dolan AM, Carnaval AC, Haywood AM, 2018. PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Sci Data* **5**: 180254.

Brown RP, Yang Z, 2010. Bayesian dating of shallow phylogenies with a relaxed clock. *Syst Biol* **59**: 119–131.

Burbrink FT, Lawson R, Slowinski JB, 2000. Molecular phylogeography of the North American rat snake *Elaphe obsoleta*: A critique of the subspecies concept. *Evolution* **54**: 2107–2118.

Burton A, 2021. Bithynian snake bombs. *Front Ecol Environ* **19**: 196.

Chiocchio A, Arntzen JW, Martínez-Solano I, de Vries W, Bisconti R et al., 2021. Reconstructing hotspots of genetic diversity from glacial refugia and subsequent dispersal in Italian common toads *Bufo bufo*. *Sci Rep* **11**:260.

De Queiroz A, Lawson R, Lemos-Espinal JA, 2002. Phylogenetic relationships of North American garter snakes (Thamnophis) based on four mitochondrial genes: How much DNA sequence is enough? *Mol Phylogenet Evol* **22**:315–329.

Drummond AJ, Ho SYW, Rawlence N, Rambaut A, 2007. *A Rough Guide to BEAST* 1.4. Available from: https://www.ccg.unam.mx/~vinuesa/tlem/docs/BEAST14_Manual_6July2007.pdf

Drummond AJ, Rambaut A, Shapiro B, Pybus OG, 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* **22**: 1185–1192.

Dufresnes C, Mazepa G, Jablonski D, Oliveira RC, Wenseleers T et al., 2019. Fifteen shades of green: the evolution of *Bufotes* toads revisited. *Mol Phylogenet Evol* **141**: 106615.

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**: 129151.

Esin NV, Esin NI, Podymov IS, Lifanchuk AV, Melnikova IV, 2019. Formation mechanisms of the Caspian transgressive seas in the Pleistocene. *Hydrosphere Ecol* **1**:13–23.

Fordham DA, Saltré F, Haythorne S, Wigley TM, Otto-Bliesner BL et al., 2017. PaleoView: A tool for generating continuous climate projections spanning the last 21000 years at regional and global scales. *Ecography* **40**:1348–1358.

Fritz U, Ayaz D, Hundsdörfer AK, Kotenko T, Guicking D et al., 2009. Mitochondrial diversity of European pond turtles *Emys orbicularis* in Anatolia and the Ponto-Caspian region: Multiple old refuges, hotspot of extant diversification and critically endangered endemics. *Org Divers Evol* **9**:100–114.

Fritz U, Corti C, Päckert M, 2012. Mitochondrial DNA sequences suggest unexpected phylogenetic position of Corso-Sardinian grass snakes *Natrix cetti* and do not support their species status, with notes on phylogeography and subspecies delineation of grass snakes. *Org Divers Evol* **12**:71–80.

Ganopolski A, Calov R, Claussen M, 2010. Simulation of the last glacial cycle with a coupled climate ice-sheet model of intermediate complexity. *Climate Past* **6**:229–244.

GBIF, 2022. GBIF Occurrence Download. https://doi.org/10.15468/dl.rd3awv, accessed on 09 March 2022.

Gerasimov IP, 1976. The main stages in the development of the relief of the Turan plains in recent geological time. In: Gerasimov IP, editor. *New Pathways in Geomorphology and Paleogeography*. Moscow: Nauka. 20–41 (in Russian).

Göçmen B, Mebert K, 2011. The rediscovery of *Natrix tessellata* on Cyprus. *Mertensiella* **18**:383–387.

Gruschwitz M, Lenz S, Mebert K, Lanka V, 1999. *Natrix tessellata (Laurenti, 1768)—Würfelnatter*. In: Böhme W, editor. *Handbuch der Reptilien und Amphibien Europas, Band 3/IIA., Schlangen (Serpentes) II*. Wiesbaden: Aula-Verlag, 581–644.

Guicking D, Joger U, 2011. A range-wide molecular phylogeography of *Natrix tessellata*. *Mertensiella* **18**:1–10.

Guicking D, Joger U, Wink M, 2002. Molecular phylogeography of the viperine snake *Natrix maura* and the dice snake *Natrix tessellata*: First results. *Biota* **3**:49–59.

Guicking D, Joger U, Wink M, 2009. Cryptic diversity in a Eurasian water snake (*Natrix tessellata*, Serpentes: Colubridae): Evidence from mitochondrial sequence data and nuclear ISSR-PCR fingerprinting. *Org Divers Evol* **9**:201–214.

Guicking D, Lawson R, Joger U, Wink M, 2006. Evolution and phylogeny of the genus *Natrix* (Serpentes: Colubridae). *Biol J Linn Soc* **87**:127–143.

Guo X, Dai X, Dali C, Papenfuss TJ, Ananjeva NB et al., 2011. Phylogeny and divergence times of some racerunner lizards (Lacertidae: Eremias) inferred from mitochondrial 16S rRNA gene segments. *Mol Phylogenet Evol* **61**:400–412.

Guo Z, Peng S, Hao Q, Biscaye PE, An Z et al., 2004. Late Miocene-Pliocene development of Asian aridification as recorded in the Red-Earth formation in northern China. *Glob Planet Change* **41**:135–145.

Hall TA, 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**:95–98.

Hanley JA, McNeil BJ, 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**:29–36.

Heller R, Chikhi L, Siegismund HR, 2013. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS One* **8**:e62992.

Hofmann S, Kraus S, Dorge T, Nothnagel M, Fritzsche P et al., 2014. Effects of Pleistocene climatic fluctuations on the phylogeography, demography and population structure of a high-elevation snake species *Thermophis baileyi* on the Tibetan Plateau. *J Biogeogr* **41**:2162–2172.

Hofmann S, Mebert K, Schulz K-D, Helfenberger N, Göçmen B et al., 2018. A new subspecies of *Zamenis hohenackeri* (Strauch, 1873) (Serpentes: Colubridae) based on morphological and molecular data. *Zootaxa* **137**:4471–4153.

Jablonski D, Nagy ZT, Avcı A, Olgun K, Kukushkin OV et al., 2019. Cryptic diversity in the smooth snake *Coronella austriaca*. *Amphib-Reptilia* **40**:179–192.

Jablonski D, Ribeiro-Júnior MA, Meiri S, Maza E, Kukushkin OV et al., 2021. Morphological and genetic differentiation in the anguid lizard *Pseudopus apodus* supports the existence of an endemic subspecies in the Levant. *Vertebr Zool* **71**:175–200.

Karger DN, Nobis MP, Normand S, Graham CH, Zimmermann NE, 2023. CHELSA-TraCE21k-high-resolution (1 km) downscaled transient temperature and precipitation data since the Last Glacial Maximum. *Clim Past* **19**:439–456.

Kass RE, Raftery AE, 1995. Bayes factors. *J Am Stat Assoc* **90**:773–795.

Kindler C, Böhme W, Corti C, Gvoždík V, Jablonski D et al., 2013. Mitochondrial phylogeography, contact zones and taxonomy of grass snakes (*Natrix natrix, N. megalocephala*). *Zool Scripta* **42**:458–472.

Kindler C, Chèvre M, Ursenbacher S, Böhme W, Hille A et al., 2017. Hybridization patterns in two contact zones of grass snakes reveal a new Central European snake species. *Sci Reports* **7**:7378.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S et al., 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc Nat Acad Sci U S A* **86**:6196–6200.

Kotenko TI, Shaitan SV, Starkov VG, Zinenko OI, 2011. The northern range limit of the dice snake *Natrix tessellata* in Ukraine and the Don River basin in Russia. *Mertensiella* **18**:311–324.

Krijgsman W, Tesakov A, Yanina T, Lazarev S, Danukalova G et al., 2019. Quaternary time scales for the Pontocaspian domain: Interbasinal connectivity and faunal evolution. *Earth-Sci Rev* **188**:1–40.

Kyriazi P, Kornilios P, Nagy ZT, Poulakakis N, Kumlutaş Y et al., 2013. Comparative phylogeography reveals distinct colonization patterns of Cretan snakes. *J Biogeogr* **40**:1143–1155.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B, 2017. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol* **34**:772–773.

Lartillot N, Philippe H, 2006. Computing Bayes factors using thermodynamic integration. *Syst Biol* **55**:195–207.

Lazarev S, Kuiper KF, Oms O, Bukhsianidze M, Vasilyan D et al., 2021. Five-fold expansion of the Caspian Sea in the late Pliocene: New and revised magnetostratigraphic and 40Ar/39Ar age constraints on the Akchagylian stage. *Glob Planet Change* **206**:103624.

Leigh JW, Bryant D, 2015. PopART: full-feature software for haplotype network construction. *Methods Ecol Evol* **6**:1110–1116.

Lemey P, Rambaut A, Welch JJ, Suchard MA, 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Mol Biol Evol* **27**:1877–1885.

Litvinov N, Bakiev A, Mebert K, 2011. Thermobiology and microclimate of the dice snake at its northern range limit in Russia. *Mertensiella* **18**:330–336.

Liu Y, Mebert K, Shi L, 2011. Notes on distribution and morphology of the dice snake *Natrix tessellata* in China. *Mertensiella* **18**:430–436.

Marosi B, Zinenko OI, Ghira IV, Crnobrnja-Isailović J, Lymberakis P et al., 2012. Molecular data confirm recent fluctuations of northern border of dice snake *Natrix tessellata* range in Eastern Europe. *North-Western J Zool* **8**:374–377.

Mebert K, 2011. *The Dice Snake Natrix tessellata: Biology, Distribution and Conservation of a Palaearctic Species*. Mertensiella, Vol. **18**. Rheinbach: DGHT.

Mebert K, Masroor R, 2013. Dice snakes in western Himalaya: Insights into regional expansion routes of *Natrix tessellata* after its rediscovery in Pakistan. *Salamandra* **49**:229–233.

Mebert K, Masroor R, Chaudhry MJI, 2013. The dice snake *Natrix tessellata* (Serpentes: Colubridae) in Pakistan: Analysis of its range limited to few valleys in the western Karakoram. *Pakistan J Zool* **45**:395–410.

Melville J, Hale J, Mantziou G, Ananjeva NB, Milto K et al., 2009. Historical biogeography, phylogenetic relationships and intraspecific diversity of agamid lizards in the Central Asian deserts of Kazakhstan and Uzbekistan. *Mol Phylogenet Evol* **53**:99–112.

Naidina OD, Richards K, 2020. The Akchagylian stage (late Pliocene-early Pleistocene) in the North Caspian. *Quater Internatonal* **540**:22–37.

Pauwels OSG, Kadeyeva M, Kovshar V, Sakharbayev A, Sarayev FA et al., 2020. Colonization of artificial islands in the Kazakh sector of the Caspian Sea by the aquatic snake *Natrix tessellata* (Squamata: Natricidae). *Bull Chicago Herpetol Soc* **55**:133–140.

Phillips SJ, Anderson RP, Schapire RE, 2006. Maximum entropy modeling of species geographic distributions. *Ecol Model* **190**:231–259.

Phillips SJ, Dudık M, Schapire RE, 2004. A maximum entropy approach to species distribution modelling. In: Proc of the 21st Intern Confer on Machine Learning, New York, 4–8 July, 655–662.

Pickarski N, Kwiecien O, Langgut D, Litt T, 2015. Abrupt climate and vegetation variability of eastern Anatolia during the last glacial. *Clim Past* **11**:1491–1505.

Popov SV, Shcherba IG, Ilyina LB, Nevesskaya LA, Paramonova NP et al., 2006. Late Miocene to Pliocene palaeogeography of the Paratethys and its relation to the Mediterranean. *Palaeogeogr Palaeoclim Palaeoecol* **238**:91–106.

Posada D, Crandall KA, 2001. Intraspecific gene genealogies: Trees grafting into networks. *Trends Ecol Evol* **16**:37–45.

Poulakakis N, Kapli P, Lymberakis P, Trichas A, Vardinoyiannis K et al., 2015. A review of phylogeographic analyses of animal taxa from the Aegean and surrounding regions. *J Zool Syst Evol Res* **53**:18–32.

Préau C, Trochet A, Bertrand R, Isselin-Nondedeu F, 2018. Modeling potential distributions of three European amphibian species comparing ENFA and MaxEnt. *Herpetol Conserv Biol* **13**:91–104.

Rajabizadeh M, Javanmardi S, Rastegar-Pouyani N, Karamiani R, Yousefi M et al., 2011. Geographic variation, distribution, and habitat of *Natrix tessellata* in Iran. *Mertensiella* **18**:414–429.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA, 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* **67**:901–904.

Rastegar-Pouyani E, Ebrahimipour F, Hosseinian S, 2017. Genetic variability and differentiation among the populations of Dice snake *Natrix tessellata* (Sepentes, Colubridae) in the Iranian Plateau. *Biochem Syst Ecol* **72**:23–28.

Ratnikov VY, 2009. Fossil remains of modern amphibian and reptile species as the material for studying the history of their distribution. *Proc Res Inst Geol Voronezh State Univ*, Vol. **59**. Voronezh: Publishing House of Voronezh State University. 91 p. (in Russian).

Ratnikov VY, Mebert K, 2011. Fossil remains of *Natrix tessellata* from the late cenozoic deposits of the East European Plain. *Mertensiella* **18**:337–342.

Ritchie AM, Lo N, Ho SY, 2017. The impact of the tree prior on molecular dating of data sets containing a mixture of inter-and intraspecies sampling. *Syst Biol* **66**:413–425.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F et al., 2011. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform* **12**:77.

Ronquist F, Teslenko M, Van der Mark P, Ayres DL, Darling A et al., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**:539–542.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P et al., 2017. DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Mol Biol Evol* **34**:3299–3302.

Sindaco, R, Jeremcenko VK, 2008. *The Reptiles of the Western Palearctic. 1. Annotated Checklist and Distributional Atlas of the Turtles, Crocodiles, Amphisbaenians and Lizards of Europe, North Africa, Middle East and Central Asia*. Latina: Edizioni Belvedere.

Sindaco R, Venchi A, Grieco C, 2013. *The Reptiles of the Western Palearctic 2. Annotated checklist and distributional atlas of the snakes of Europe, North Africa, the Middle East and Central Asia, with an update to the vol. 1*. Latina: Edizioni Belvedere, Soc Herpetol Italica, Via Adige.

Šmíd J, Aghová T, Velenská D, Moravec J, Balej P et al., 2021. Quaternary range dynamics and taxonomy of the Mediterranean collared dwarf racer *Platyceps collaris* (Squamata: Colubridae). *Zool J Linn Soc* **193**:655–672.

Solovyeva EN, Lebedev VS, Dunayev EA, Nazarov RA, Bannikova AA et al., 2018. Cenozoic aridization in Central Eurasia shaped diversification of toad-headed agamas (*Phrynocephalus;* Agamidae, Reptilia). *PeerJ* **6**:e4543.

Stamatakis A, 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312–1313.

Stratakis M, Koutmanis I, Ilgaz C, Jablonski D, Kukushkin OV et al., 2022. Evolutionary divergence of the smooth snake (Serpentes, Colubridae): The role of the Balkans and Anatolia. *Zool Scripta* **51**:310–329.

Svitoch AA, 2013. The Pleistocene Manych straits: their structure, evolution and role in the Ponto-Caspian basin development. *Quater International* **302**:101e109.

Tuniyev BS, 1995. On the Mediterranean influence on the formation of herpetofauna of the Caucasian Isthmus and its main xerophylous refugia. *Russ J Herpetol* **2**:95–119.

Tuniyev B, Tuniyev S, Kirschey T, Mebert K, 2011. Notes on the dice snake *Natrix tessellata* from the Caucasian Isthmus. *Mertensiella* **18**:343–356.

Van Baak CGC, Krijgsman W, Magyar I, Sztanó O, Golovina LA et al., 2017. Paratethys response to the Messinian salinity crisis. *Earth-Sci Rev* **172**:193–223.

Wegwerth A, GanopolskiBard E, Lamy F, Arz HWA, Ménot G et al., 2015. Black Sea temperature response to glacial millennial-scale climate variability. *Geophys Res Lett* **42**:8147–8154.

Xie W, Lewis PO, Fan Y, Kuo L, Chen M-H, 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst Biol* **60**:150–160.

Yakovleva ID, 1964. *Reptiles of Kirgizia*. Frunze: Ilim (in Russian).

Yakovleva TI, Bakiev AG, 2010. Changes in the fauna and range boundaries of snakes in the Volga basin in the Late Cenozoic. In: Rozenberg GS, Saksonov SV, editors. *Theoretical Problems of Ecology and Evolution. Theory of Home Ranges: Species, Communities, Ecosystems (5th Lyubishchev readings)*. Togliatti: Cassandra, 221–231 (in Russian).

Zhang YJ, Stöck M, Zhang P, Wang XL, Zhou H et al., 2008. Phylogeography of a widespread terrestrial vertebrate in a barely studied Palearctic region: Green toads (*Bufo viridis* subgroup) indicate glacial refugia in Eastern Central Asia. *Genetica* **134**:353–365.

Zhou L, Liang T, Shi L, 2019. Amphibian and reptilian chorotypes in the arid land of Central Asia and their determinants. *Sci Reports* **9**:9453.

Zinenko O, Stümpel N, Mazanaeva L, Bakiev A, Shiryaev K et al., 2015. Mitochondrial phylogeny shows multiple independent ecological transition and northern dispersion despite of Pleistocene glaciations in meadow and steppe vipers (*Vipera ursinii* and *Vipera renardi*). *Mol Phylogenet Evol* **84**:85–100.