



It takes two to tango – Phylogeography, taxonomy and hybridization in grass snakes and dice snakes (Serpentes: Natricidae: *Natrix natrix*, *N. tessellata*)

Marika Asztalos¹, Dinçer Ayaz², Yusuf Bayrakçı², Murat Afsar³, Cemal Varol Tok⁴, Carolin Kindler¹, Daniel Jablonski⁵, Uwe Fritz¹

¹ Museum of Zoology (Museum für Tierkunde), Senckenberg Dresden, A. B. Meyer Building, 01109 Dresden, Germany

² Zoology Section, Department of Biology, Faculty of Science, Ege University, İzmir, Turkey

³ Department of Biology, Faculty of Arts & Sciences, Celal Bayar University, 45140, Yunusmre, Manisa, Turkey

⁴ Department of Biology, Zoology Section, Faculty of Arts and Sciences, Çanakkale Onsekiz Mart University, Terzioğlu Campus, Çanakkale, Turkey

⁵ Department of Zoology, Comenius University in Bratislava, Ilkovičova 6, Mlynská dolina, 842 15 Bratislava, Slovakia

<http://zoobank.org/08F341B3-0DE5-4902-9C12-9265A3644ADE>

Corresponding author: Uwe Fritz (uwe.fritz@senckenberg.de)

Academic editor Ralf Britz | Received 12 October 2021 | Accepted 23 November 2021 | Published 7 December 2021

Citation: Asztalos M, Ayaz D, Bayrakçı Y, Afsar M, Tok CV, Kindler C, Jablonski D, Fritz U (2021) It takes two to tango – Phylogeography, taxonomy and hybridization in grass snakes and dice snakes (Serpentes: Natricidae: *Natrix natrix*, *N. tessellata*). Vertebrate Zoology 71 813–834. <https://doi.org/10.3897/vz.71.e76453>

Abstract

Using two mitochondrial DNA fragments and 13 microsatellite loci, we examined the phylogeographic structure and taxonomy of two codistributed snake species (*Natrix natrix*, *N. tessellata*) in their eastern distribution area, with a focus on Turkey. We found evidence for frequent interspecific hybridization, previously thought to be extremely rare, and for backcrosses. This underscores that closely related sympatric species should be studied together because otherwise the signal of hybridization will be missed. Furthermore, the phylogeographic patterns of the two species show many parallels, suggestive of a shared biogeographic history. In general, the phylogeographies follow the paradigm of southern richness to northern purity, but the dice snake has some additional lineages in the south and east in regions where grass snakes do not occur. For both species, the Balkan Peninsula and the Caucasus region served as glacial refugia, with several mitochondrial lineages occurring in close proximity. Our results show that the mitochondrial divergences in both species match nuclear genomic differentiation. Yet, in the former glacial refugia of grass snakes there are fewer nuclear clusters than mitochondrial lineages, suggesting that Holocene range expansions transformed the glacial hotspots in melting pots where only the mitochondrial lineages persisted, bearing witness of former diversity. On the other hand, the deep mitochondrial divergences in *N. tessellata* across its entire range indicate that more than one species could be involved, even though lacking microsatellite data outside of Turkey prevent firm conclusions. On the contrary, our microsatellite and mitochondrial data corroborate that *N. megaloccephala* is invalid and not differentiated from sympatric populations of *N. natrix*. For Cypriot grass snakes, our analyses yielded conflicting results. A critical assessment of the available evidence suggests that *N. natrix* is a genetically impoverished recent invader on Cyprus and taxonomically not distinct from a subspecies also occurring in western Anatolia and the southern Balkans. Based on combined mitochondrial and nuclear genomic evidence we propose that for grass snakes the following subspecies should be recognized in our study region: (1) *Natrix natrix vulgaris* Laurenti, 1768, southeastern Central Europe and northern Balkans; (2) *Natrix natrix moreotica* (Bedriaga, 1882), southern Balkans, western Anatolia, and Cyprus; and (3) *Natrix natrix scutata* (Pallas, 1771), eastern Anatolia, Caucasus region, Iran, northeastern distribution range (from eastern Poland and Finland to Kazakhstan and the Lake Baikal region). Thus, *Natrix natrix cyprica* (Hecht, 1930) becomes a junior synonym of *N. n. moreotica* and *Natrix natrix persa* (Pallas, 1814) becomes a junior synonym of *N. n. scutata*. Due to insufficient material, we could not resolve the status of *Natrix natrix syriaca* (Hecht, 1930) from the Gulf of İskenderun, southeastern Turkey.

Keywords

Biogeography, hybridization, *Natrix natrix*, *Natrix megalcephala*, *Natrix tessellata*, phylogeography, subspecies, taxonomy, Turkey

Introduction

The genus *Natrix* Laurenti, 1768 comprises five widely distributed Palearctic snake species (Speybroeck et al. 2020). Among them, the viperine snake *Natrix maura* (Linnaeus, 1758) and the dice snake *Natrix tessellata* (Laurenti, 1768) are true water snakes (Gruschwitz et al. 1999; Schätti 1999) and have been regarded as distinct species since their description in the mid-18th century. In contrast, the red-eyed grass snake *Natrix astreptophora* (Seoane, 1884), the barred grass snake *Natrix helvetica* (Lacépède, 1789), and the common grass snake *Natrix natrix* (Linnaeus, 1758) have been regarded as conspecific for a long time. Compared to *N. maura* and *N. tessellata*, the grass snakes are less bound to aquatic habitats, even though they frequently occur near water bodies, with *N. astreptophora* having the least aquatic lifestyle (Kabisch 1999; Kreiner 2007; Geniez 2015; Pokrant et al. 2017).

Natrix maura is a Western Mediterranean and Western European species, ranging from northwestern Africa and the Iberian Peninsula to central and southern France and northwestern Italy. Introduced populations also occur on the Balearic Islands and Sardinia (Schätti 1999; Guicking et al. 2008; Geniez 2015). Further east, the ecologically and morphologically similar *N. tessellata* replaces *N. maura*. *Natrix tessellata* occurs across most of mainland Italy and ranges from the Czech Republic and southern and eastern Austria eastward through the south of Eastern Europe to northwestern China (Xinjiang); a few isolated Central European populations (Germany, Switzerland) are witness of a formerly wider Holocene distribution in the northwest. Along the Levantine coast, *N. tessellata* is distributed southward to Israel and northeastern Egypt (Bannikov et al. 1977; Zhao and Adler 1993; Gruschwitz et al. 1999; Geniez 2015; Speybroeck et al. 2016). While the ranges of *N. maura* and *N. tessellata* are largely mutually exclusive and only overlap in northwestern Italy, these two water snakes occur in wide sympatry with the three grass snake species (*N. astreptophora*, *N. helvetica*, *N. natrix* sensu stricto).

Until a few years ago, the three grass snake species were lumped together in one polytypic species, *N. natrix* sensu lato (Mertens and Wermuth 1960; Kabisch 1999; Kreiner 2007). However, recent studies using mtDNA sequences and microsatellite loci revealed only limited gene flow across narrow geographic contact zones and led to their recognition as three distinct species (Pokrant et al. 2016; Kindler et al. 2017; Schultze et al. 2019, 2020; Asztalos et al. 2020, 2021a; Speybroeck et al. 2020). *Natrix astreptophora* occurs in the northern Maghreb region, the Iberian Peninsula and southwestern France (Mertens and Wermuth 1960; Kabisch 1999). In southwestern France

lies a narrow contact zone of this species and *N. helvetica*, where limited bidirectional gene flow occurs (Pokrant et al. 2016; Asztalos et al. 2020). This contact zone is a bimodal hybrid zone (see Jiggins and Mallet 2000), with pure parental genotypes and hybrid genotypes occurring together or in close proximity. *Natrix helvetica* is distributed in France, Great Britain, the Benelux countries, western Germany, Switzerland, Italy and across the Alps to southern Bavaria and Tyrol (Geniez 2015; Kindler et al. 2017; Schultze et al. 2020; Asztalos et al. 2021a). The third grass snake species is *Natrix natrix* sensu stricto, which occurs from the Rhine region in the west to Lake Baikal in the east. Its vast distribution range also includes Fennoscandia, northeastern Italy, the Balkan Peninsula and Anatolia, the Transcaucasus, and northern Iran (Geniez 2015; Kindler et al. 2017; Fritz and Schmidler 2020; Schultze et al. 2020). Three narrow bimodal hybrid zones between *N. helvetica* and *N. natrix* have been described. One is located in the Rhine region, where gene flow is mainly unidirectional from *N. h. helvetica* into *N. natrix* (Kindler et al. 2017; Schultze et al. 2019). In the two other hybrid zones another subspecies of the barred grass snake is involved, *N. h. sicula* (Cuvier, 1829). It hybridizes with *N. natrix* in northeastern Italy (Schultze et al. 2020) and, beyond the Alps, in southern Bavaria and Tyrol (Asztalos et al. 2021a).

Each of the three grass snake species is polytypic. Recently, Fritz and Schmidler (2020) reconciled traditional morphology-based subspecies with genetically differentiated units. These authors presented a classification scheme recognizing geographically vicariant, evolutionarily significant genetic clusters as distinct subspecies. Only for the Balkan Peninsula, Anatolia and adjacent regions including the Caucasus and Transcaucasus, the intraspecific variation of *N. natrix* could not be assessed because no nuclear-genomic evidence was then available. While intraspecific variation in grass snakes is now, with the mentioned notable exception, well understood, for *N. maura* and *N. tessellata* the situation is strikingly different. Both species contain genetically deeply divergent parapatric lineages (Guicking et al. 2008, 2009; Guicking and Joger 2011) suggestive of taxonomic divergence. Yet, in contrast to grass snakes, traditionally no subspecies have been recognized for *N. maura* (Mertens and Wermuth 1960; Schätti 1999; Geniez 2015; Speybroeck et al. 2016). For *N. tessellata*, only the micro-endemic subspecies *Natrix tessellata heinrothi* (Hecht, 1930) from the small island of Serpilor in the Black Sea has been recognized by some authors. However, its validity has been disputed (see Gruschwitz et al. 1999; Geniez 2015).

In the present study we aim to resolve the taxonomic and genetic differentiation of grass snakes (*N. natrix* sensu stricto) in the Balkan Peninsula, Anatolia, the Caucasus region and Iran. In addition, we scrutinize whether the phylogeographic and taxonomic differentiation of the dice snake (*N. tessellata*) parallels that of *N. natrix* in our study region, and we examine, for the first time, whether and, if so, to which extent, natural hybridization occurs between these two sympatric species. For doing so, we apply an established toolbox consisting of combined evidence from mitochondrial DNA (mtDNA) sequences and 13 nuclear genomic microsatellite loci (Kindler et al. 2013, 2017, 2018a; Kindler and Fritz 2018; Pokrant et al. 2016; Aszталos et al. 2020, 2021a, b; Schultze et al. 2019, 2020) to examine 683 grass snakes and 332 dice snakes from our immediate study region and beyond. We also reassess the status of *N. megalcephala* Orlov and Tuniyev, 1987 based on nuclear genomic microsatellite loci. This western Caucasian species was described based on slight morphological differences to putatively syntopic *N. natrix* (see Orlov and Tuniyev 1987, 1992). However, in the face of inconsistent morphological characters and the lack of any distinctiveness in mtDNA, the validity of *N. megalcephala* was soon challenged and the name was synonymized with *N. natrix*. *Natrix megalcephala* is currently identified as a junior synonym of the subspecies *N. n. scutata* (Pallas, 1771) (Kindler et al. 2013; see the review in Fritz and Schmidler 2020).

Materials and Methods

Sampling and laboratory procedures

Data for 683 grass snakes (*Natrix natrix*) and 332 dice snakes (*N. tessellata*) were analyzed for the present study. A total of 181 tissue samples (115 *N. natrix*, 66 *N. tessellata*) were newly processed and microsatellite data were generated for 15 additional samples of *N. natrix* with previously published mtDNA sequences (Kindler et al. 2013, 2017). The remaining data originate from earlier studies (Guicking et al. 2006, 2009; Kindler et al. 2013, 2017; Kyriazi et al. 2013; Aszталos et al. 2021a). Among our material were five samples of big-headed grass snakes from the putative distribution range of *N. megalcephala* that were, according to their morphology, tentatively identified with this nominal species (Table S1).

Two marker systems (mtDNA, microsatellite loci) were applied, which were successfully used in our previous studies on the phylogeography and hybridization of grass snakes (Kindler et al. 2013, 2017, 2018a, b; Pokrant et al. 2016; Kindler and Fritz 2018; Schultze et al. 2019, 2020; Aszталos et al. 2020, 2021a, b; Ahnelt et al. 2021).

Laboratory procedures followed Kindler et al. (2013) and Pokrant et al. (2016). Samples collected within the frame of TÜBİTAK project 116Z359 in Turkey were processed in the Biodiversity Research Laboratory

of Ege University, İzmir. The remaining new material was studied in the molecular laboratory of Senckenberg Dresden.

Two mtDNA fragments were sequenced, the cytochrome *b* (*cyt b*) gene (1,117 bp) and the partial ND4 gene plus adjacent DNA coding for tRNAs (tRNA-His, tRNA-Ser, tRNA-Leu; 866 bp). For some samples the obtained sequences were shorter and for a few samples, only one mtDNA fragment could be sequenced (Table S1).

In addition to mtDNA sequencing, the samples were also genotyped at the same 13 polymorphic microsatellite loci as in our previous studies. For *N. tessellata*, the presence of the microsatellite motifs was verified by sequencing the PCR products in both directions using unlabeled primers. For detailed information on microsatellite loci, primers and PCR conditions for microsatellites and mtDNA, see Tables S2 and S3, Kindler et al. (2013), and Pokrant et al. (2016). Allele lengths of material processed in İzmir were calibrated using 50 samples from the Balkans that were studied in both laboratories. For a few of the genotyped samples, no mtDNA sequences could be obtained (Table S1).

Mitochondrial DNA sequences and network analyses

Cyt b sequences could be generated for 110 samples (55 *Natrix natrix*, 55 *N. tessellata*), ND4+tRNAs for 112 samples (73 *N. natrix*, 39 *N. tessellata*). These sequences were aligned with those from previous studies (Guicking et al. 2006, 2009; Kindler et al. 2013, 2014, 2017, 2018a; Kyriazi et al. 2013; Kindler and Fritz 2018; Schultze et al. 2019, 2020; Aszталos et al. 2021a) using BIOEDIT 7.0.5.2 (Hall 1999). To identify mitochondrial haplotypes, previously characterized haplotypes of the two species (Guicking et al. 2006, 2009; Kindler et al. 2017, 2018a; Schultze et al. 2019, 2020; Ahnelt et al. 2021; Aszталos et al. 2021a, b) were added to the alignments of each mtDNA block, resulting in a 1,117-bp-long alignment of 705 *cyt b* sequences for *N. natrix* and 435 *cyt b* sequences for *N. tessellata* and an 866-bp-long alignment of 722 ND4+tRNAs sequences for *N. natrix* and 38 ND4+tRNAs sequences for *N. tessellata*. The alignment for ND4+tRNAs of *N. tessellata* included only the sequences generated for the present study, whereas the other alignments also contained all previously published homologous sequence data. For each alignment, parsimony networks were drawn using TCS 1.21 (Clement et al. 2000), with gaps coded as fifth character state. Using the default 95% connection limit, unconnected haplotype clusters were obtained for the different genetic lineages within each species, which is why the connection limit was arbitrarily set to 150 steps. Haplotype terminology for *N. natrix* follows Kindler et al. (2013) and for the *cyt b* of *N. tessellata*, Guicking et al. (2009). For the ND4+tRNAs fragment, which was not studied by Guicking et al. (2009), an analogous system is introduced here in that each haplotype is consecutively numbered and this number follows the code letter for the haplotype cluster (e.g., E1, E2, etc.).

For each of the two species and for individual lineages within each species, uncorrected p distances (averages) were computed based on the haplotypes of each mtDNA fragment using MEGA 10.2.3 (Kumar et al. 2018) and the pairwise deletion option.

Microsatellite cluster analyses, inferring hybrid status and PCAs

Our microsatellite data were subjected to unsupervised Bayesian cluster analyses using STRUCTURE 2.3.4 (Pritchard et al. 2000; Falush et al. 2003). MICRO-CHECKER 2.2.3 (van Oosterhout et al. 2004) revealed no evidence for null alleles, necessitating no corrections for null alleles. STRUCTURE was run using the admixture model and correlated allele frequencies. All calculations were repeated 10 times for each K ranging from 1 to 10, using a Monte Carlo Markov chain of 1,000,000 generations and a burn-in of 250,000. To infer the most likely number of clusters (K), STRUCTURE HARVESTER (Earl and vonHoldt 2012) was used, a software that implements the ΔK method of Evanno et al. (2005). Population structuring and individual admixture were visualized using the R package POPHELPER 2.2.9 (Francis 2017). Since STRUCTURE is known to detect only the uppermost level of population differentiation (Evanno et al. 2005), different hierarchical runs were executed: A first calculation included all data from 639 *Natrix natrix* and 66 *N. tessellata*. Then, data for all individuals with evidence for interspecific hybridization were removed and a pruned dataset for each species alone was processed. Within each species, further analyses were run for lower hierarchical levels to examine for finer substructuring. For each of these analyses, data for snakes with evidence for admixture between the relevant interspecific clusters were excluded (Table S1).

To infer whether individual genotypes in STRUCTURE analyses represent hybrids or pure individuals, thresholds for hybrid identity were determined using HYBRIDLAB 1.0 (Nielsen et al. 2006). For doing so, from the first STRUCTURE run 25 pure individuals each of *N. natrix* and of *N. tessellata* were selected as parental genotypes and the cluster membership coefficients (Q values) from the four clusters corresponding to *N. natrix* were summed up and opposed to the values of *N. tessellata* (for the rationale, see under Results and Discussion). Then, 25 genotypes of each hybrid class (F_1 , F_2 , two backcrosses) were modeled and used for inferring the thresholds for pure genotypes of each parental species (Table S4). The same approach was used for finding thresholds for pure clusters within *N. natrix* (Table S5). Inferred hybrid thresholds are summarized in the Supplementary Tables S4 and S5. For the lowermost hierarchical analysis for *N. natrix* and for *N. tessellata*, the ΔK method suggested the presence of three and four distinct clusters, respectively. Since HYBRIDLAB cannot infer hybrid thresholds for more than two clusters, an arbitrary threshold was applied here and genotypes with proportions of cluster membership below 80% were treated as admixed (Randi 2008).

STRUCTURE analyses are based on population-genetic presumptions (Hardy-Weinberg equilibrium, linkage equilibrium; Pritchard et al. 2000) and therefore prone to bias from uneven sample sizes (Puechmaille 2016). To cross-examine STRUCTURE results for potential artifacts, additional Principal Component Analyses (PCAs) were run using the R package ADEGENET 2.1.1 (Jombart 2008). PCAs are independent from any population-genetic presumptions and exclusively based on genetic information, rendering them less sensitive to different sample sizes. PCAs were computed for the same datasets as used for STRUCTURE analyses, and the individual results were depicted either according to their morphological or mitochondrial and nuclear genomic identity and compared to our STRUCTURE results.

Inference of population-genetic parameters

For selected populations, allelic richness was calculated using FSTAT 2.9.4 (Goudet 1995). The number of alleles (n_A), the average number of alleles per locus ($n_{\bar{A}}$), average observed and expected heterozygosity (H_O and H_E) and the inbreeding coefficient (F_{IS}) were inferred using ARLEQUIN 3.5.2.2 (Excoffier and Lischer 2010).

Results

Mitochondrial DNA

In parsimony network analyses (cyt *b*, ND4+tRNAs), each mitochondrial lineage of *Natrix natrix* and *N. tessellata* revealed by phylogenetic analyses (Guicking et al. 2009; Guicking and Joger 2011; Kindler et al. 2013) corresponded to a distinct haplotype cluster.

Haplotype networks for grass snakes (*Natrix natrix*)

The network for cyt *b* of *Natrix natrix* contained several loops (Fig. 1). The haplotype cluster corresponding to lineage 7 was located in the center of the network and connected to the haplotype clusters of lineages 1, 2, 3, 5, 6, and 8. Lineage 7 comprised 31 haplotypes that differed by a maximum of 16 mutations. The haplotype cluster of lineage 7 was connected by a minimum of 27 mutation steps to the haplotypes of lineage 2 and by a minimum of 29 mutations to those of lineage 1. The haplotypes of lineages 1 and 2 differed from one another by a minimum of 12 mutations. Among the five haplotypes of lineage 2 a maximum of 15 mutations occurred; among the six haplotypes of lineage 1, a maximum of eight mutations. Haplotypes of lineage 5 were connected to haplotypes of lineages 1, 2, and 7 by a minimum of 58, 56, and 39 mutations, respectively. Lineage 5 comprised 25 haplotypes

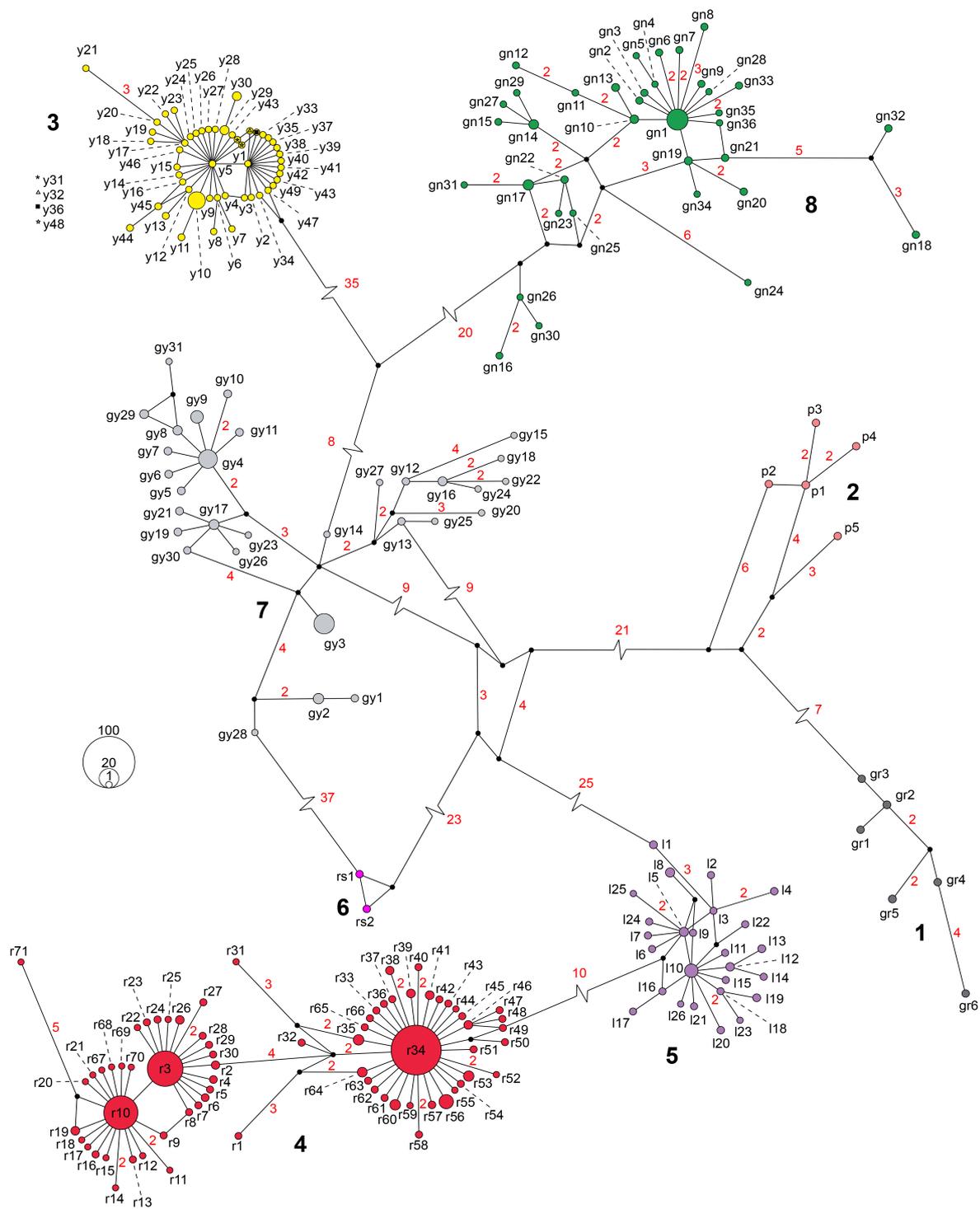


Figure 1. Parsimony network for 705 *cyt b* sequences of *Natrix natrix*. Symbol sizes reflect haplotype frequencies. Small black circles are missing node haplotypes; each line connecting two haplotypes corresponds to one mutation step, if not otherwise indicated by red numbers along lines. Haplotype colors correspond to lineages.

that differed by a maximum of seven mutations. Haplotypes of lineage 4 were only connected to lineage 5 by a minimum of 11 mutations. The 71 haplotypes of lineage 4 differed by a maximum of 18 mutations. Haplotypes of lineage 6 were connected to haplotypes of lineages 5 and 7 and differed by a minimum of 50 mutations from

haplotypes of lineage 5 and by a minimum of 37 mutations from those of lineage 7. Lineage 6 comprised two haplotypes that differed by one mutation. Haplotypes of lineage 7 were also connected with haplotypes of lineages 3 and 8. Haplotypes of lineage 3 differed by a minimum of 44 mutations from haplotypes of lineage 7 and

by a minimum of 57 mutations from those of lineage 8. Among the 49 haplotypes of lineage 3 occurred up to seven mutations. Haplotypes of lineage 8 were connected by a minimum of 21 mutations to haplotypes of lineage 7. Lineage 8 consisted of 36 haplotypes that differed by a maximum of 19 mutations.

The following *cyt b* haplotypes were newly identified in the present study: gn22–gn27, gy15–gy31, and l23 (European Nucleotide Archive accession numbers OU862648–OU862671).

The network for ND4+tRNAs sequences of *N. natrix* (Fig. S1) resembled the *cyt b* network, but with fewer mutations, and the clusters of lineages 4 and 5 were not clearly distinct. Again, the haplotype cluster of lineage 7 was located in the center of the network, with direct connections to haplotypes of lineages 1, 3, 5, 6, and 8. The 30 individual haplotypes of lineage 7 differed by up to 14 mutations. The haplotypes of lineage 7 were connected to haplotypes of lineage 1 by a minimum of 35 mutations. Lineage 1 comprised three haplotypes that differed by a maximum of four mutations. Haplotypes of lineage 2 were connected to haplotypes of lineage 1 by a minimum of four mutations. The four haplotypes of lineage 2 differed by a maximum of three mutations. In addition, haplotypes of lineage 7 were connected over at least 25 mutations to haplotypes of lineage 5. The latter lineage comprised 20 individual haplotypes that differed by a maximum of eight mutations. Lineage 5 was connected with haplotypes of lineage 4 by a minimum of only 2 mutations. Lineage 4 comprised 39 haplotypes that differed by a maximum of ten mutations. The only haplotype of lineage 6 was connected to lineage 7 by a minimum of ten mutations. Furthermore, haplotypes of lineage 7 were also connected to those of lineages 3 (minimum 29 mutations) and 8 (minimum 16 mutations). Haplotypes of lineages 3 and 8 were separated by a minimum of 37 mutations. The 48 haplotypes of lineage 3 and the 21 haplotypes of lineage 8 differed both by a maximum of nine mutations.

The following haplotypes for ND4+tRNAs were newly identified in the present study: gn17–gn21, gy13–gy31, and r39 (European Nucleotide Archive accession numbers OU862605–OU862629). All currently known haplotypes for *N. natrix* and their accession numbers are listed in the Supplementary Information (Table S6).

Haplotype networks for dice snakes (*Natrix tessellata*)

For *cyt b* of *Natrix tessellata* (Fig. 2), the 32 haplotypes of lineage T were located in the interior of the network of all haplotype clusters. Direct connections existed to the haplotypes of lineages A, E, G, I, and K. Among the haplotypes of lineage T occurred a maximum of 54 mutations. The haplotypes of lineage T were connected by a minimum of 46 mutations to those of lineage G and by a minimum of 43 mutations to the haplotypes of lineage I. The haplotypes of lineages G and I were separated by a minimum of 87 mutations. The 15 haplotypes of lineage G differed by a maximum of 47 mutations, while the three haplotypes of lineage I differed by a maximum of six mutations. Lineage

E comprised 66 haplotypes, which differed by a maximum of 19 mutations. The haplotypes of lineage E were connected by a minimum of 35 mutations to lineage T. The haplotypes of lineage C were connected to those of lineage E by a minimum of 26 mutations. The four haplotypes of lineage C differed by a maximum of five mutations. Lineage T was also connected to the haplotypes of lineages A and K and differed by a minimum of 19 mutations from lineage A and by a minimum of 21 mutations from lineage K. The haplotypes of lineages A and K differed by a minimum of 24 mutations from one another. Lineage K comprised 12 haplotypes that differed by a maximum of 12 mutations. Lineage A was comprised of 12 haplotypes that differed by a maximum of 11 mutations. Lineage A differed from lineages J and U by a minimum of 25 and 19 mutations, respectively. The three haplotypes of lineage J differed by a maximum of three mutations, the six haplotypes of lineage U by a maximum of 12 mutations.

The following *cyt b* haplotypes were newly identified in the present study: E52–E61 and T16–T24 (European Nucleotide Archive accession numbers OU862672–OU862690).

The second network (ND4+tRNAs) for *N. tessellata* was restricted to two lineages, E and T (Fig. S2). The haplotypes of the two lineages were connected over two pathways by a minimum of 44 mutations each. Lineage E corresponded to eight haplotypes, which differed by a maximum of six mutations. The 10 haplotypes of lineage T differed by up to 28 mutations.

All haplotypes for ND4+tRNAs were newly identified in the present study: E1–E8 and T1–T10 (European Nucleotide Archive accession numbers OU862630–OU862647). These haplotypes and all currently known *cyt b* haplotypes for *N. tessellata* and their accession numbers are listed in the Supplementary Information (Table S6).

Mean uncorrected *p* distances

Using distinct haplotypes, the average uncorrected *p* distance between dice snakes and grass snakes was for the *cyt b* gene 11.50%, and for the DNA coding for the partial ND4 gene plus tRNAs 12.10% (Table S7). For the eight mitochondrial lineages of *Natrix natrix*, divergences of 0.96–4.21% were observed for the *cyt b* gene, and for ND4+tRNAs, 0.60–5.17%. The nine mitochondrial lineages of *N. tessellata* were more divergent. In the *cyt b* gene, they differed on average by 1.34–6.15%, and for ND4+tRNAs haplotypes, a divergence of 6.93% was observed for lineages E and T (Table S8). The highest divergence values within *N. tessellata* resemble those between the three grass snake species *N. astreptophora*, *N. helvetica*, and *N. natrix* (Kindler et al. 2017, 2018a).

Geographic distribution of mitochondrial lineages

When the geographic distribution of the individual mitochondrial lineages of the two *Natrix* species is compared (Fig. 3), some similarities and some differences become

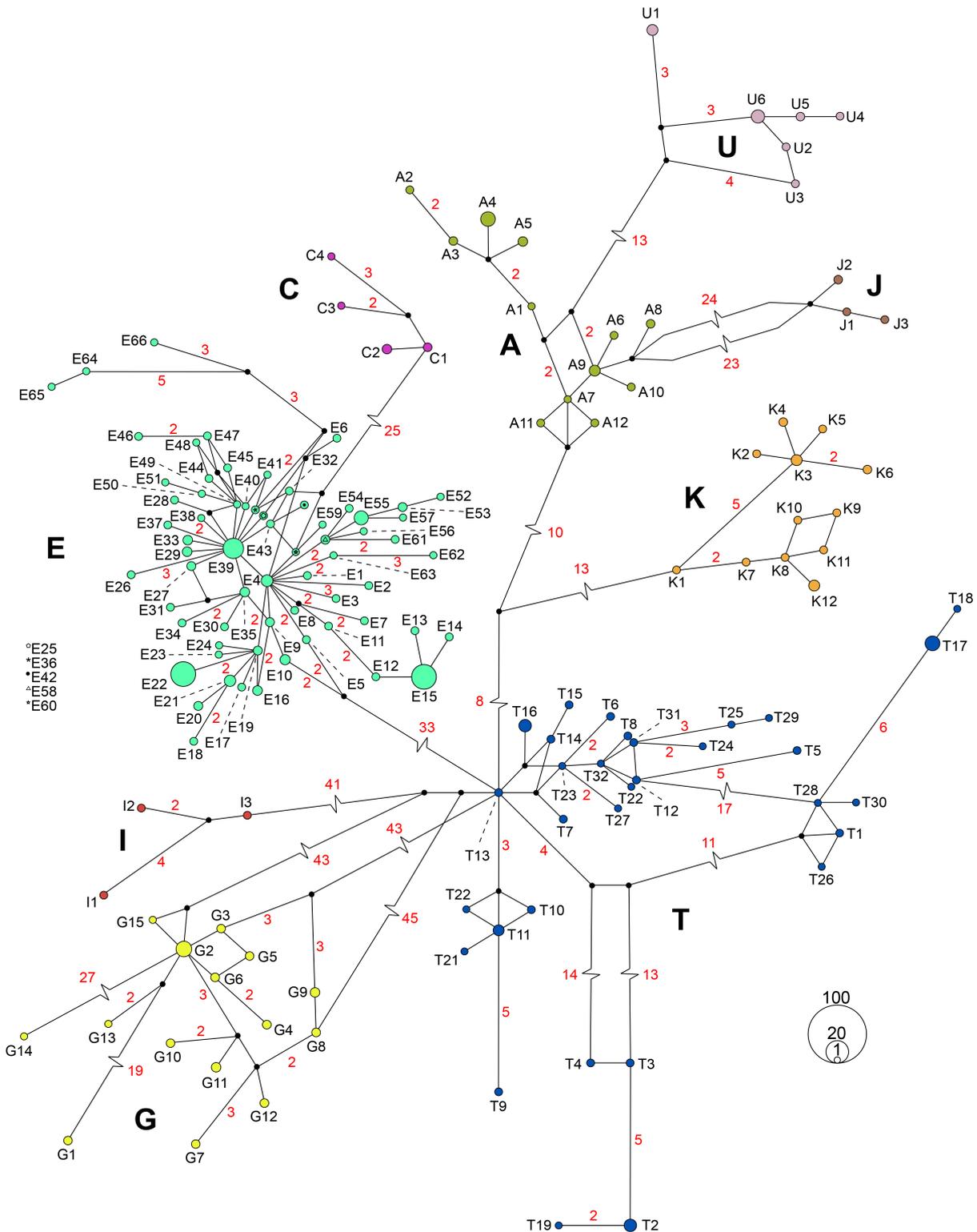


Figure 2. Parsimony network for 435 *cyt b* sequences of *Natrix tessellata*. Symbol sizes reflect haplotype frequencies. Small black circles are missing node haplotypes; each line connecting two haplotypes corresponds to one mutation step, if not otherwise indicated by red numbers along lines. Haplotype colors correspond to lineages.

apparent. In both species, there is a widely distributed Balkan lineage (lineage 4 in *N. natrix*; lineage E in *N. tessellata*). Beyond the Balkan Peninsula and beyond the scope of the present study, these lineages also encroach on regions further north and west (Guicking et al. 2009;

Guicking and Joger 2011; Kindler et al. 2013, 2017). On the Peloponnese, and in *N. natrix* also northwest of the Peloponnese, this Balkan lineage is replaced by another lineage (lineage 5 in *N. natrix*; lineage G in *N. tessellata*). However, in *N. natrix*, there are two other lineages (3 and

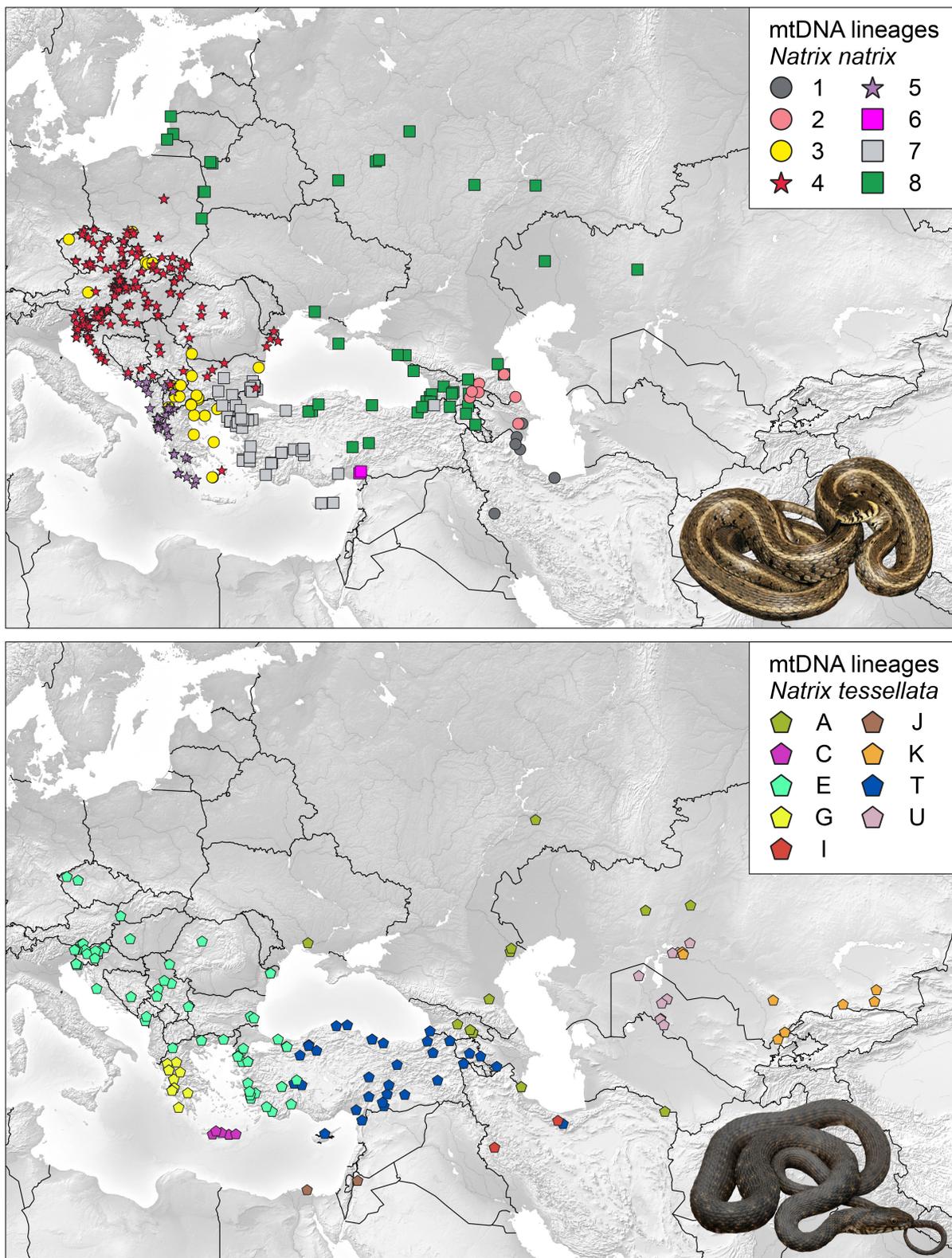


Figure 3. Sampling sites and geographic distribution of mtDNA lineages for 653 grass snakes (*Natrix natrix*; top) and 325 dice snakes (*N. tessellata*; bottom) used in this study (eastern parts of the distribution ranges). For *N. tessellata*, data outside Turkey are from Guicking et al. (2006, 2009) and Kyriazi et al. (2013). Insets: *Natrix natrix* (top; Ropotamo, Bulgaria) and *N. tessellata* (bottom; Ömeroba, Edirne, Turkey). Photos: Daniel Jablonski.

7) in the southeastern Balkan Peninsula, western Anatolia and Cyprus, whereas in *N. tessellata* the widely distributed Balkan lineage E also occupies the southeastern

Balkan Peninsula and western Anatolia. In both species there is a divide in Anatolia between a western lineage (lineage 7 in *N. natrix*; lineage E in *N. tessellata*) and an

eastern lineage (lineage 8 in *N. natrix*; lineage T in *N. tessellata*). Yet, in *N. tessellata* the eastern lineage is found on Cyprus (albeit based on only a single specimen), and in *N. natrix* the western lineage occurs on the island. Furthermore, in *N. natrix* there is a unique endemic lineage in the Gulf of İskenderun (lineage 6), whereas the range of *N. tessellata* reaches further south to northeastern Egypt, where another distinct lineage (J) occurs. In both species, several mitochondrial lineages meet in the Caucasus region¹, with a unique endemic lineage (2) in *N. natrix* in the eastern Caucasus region. In both species, a widely distributed lineage occurs north of the Black Sea and the Caspian Sea (lineage 8 in *N. natrix*; lineage A in *N. tessellata*), however, lineage 8 in *N. natrix* encroaches on eastern Anatolia and occupies there the area where lineage T in *N. tessellata* occurs. East of the Caspian Sea, two additional lineages occur in *N. tessellata* (lineages K and U), in regions where *N. natrix* is absent.

Nuclear genetic evidence

Interspecific hybridization and the status of Cypriot snakes

For each of the two species, the 13 studied microsatellite loci were highly polymorphic, with 9–32 alleles per locus for *Natrix natrix* and 3–19 alleles per locus for *N. tessellata*. Total allele numbers for *N. natrix* were 229 and for *N. tessellata*, 158 (for allele size range and individual allele numbers per locus, see Table S3).

After initial analysis of our total dataset ($n = 705$) using STRUCTURE 2.3.4, hidden structuring was examined using hierarchically nested subsets corresponding to each of the two species and genetic clusters within either species, with admixed individuals in each step removed. The respective results were then compared to the mitochondrial identity of each sample. In addition to the STRUCTURE analyses, PCAs were calculated for the same datasets to scrutinize the STRUCTURE results using an approach free of population-genetic assumptions.

For the STRUCTURE run using all 705 genotypes, the ΔK method identified $K = 2$ as the optimal number of clusters. However, a second pronounced peak occurred for $K = 5$ (Fig. S3A). Under $K = 2$, one cluster matched *Natrix natrix* and the other *N. tessellata* (Fig. 4); 555 genotypes were assigned to *N. natrix*, 77 to *N. tessellata*, and the remaining 73 were identified as interspecific hybrids. The genotypes of ‘*N. megalcephala*’ were either not dif-

ferentiated from *N. natrix* ($n = 2$) or identified as hybrids ($n = 2$).

Most of the interspecific hybrids were from Turkey, but hybrid signatures were also found in the Balkans, Transcaucasia and Russia. Sixty-six hybrids were morphologically identified as grass snakes (64 as *N. natrix*, 2 as ‘*N. megalcephala*’) and the remaining seven hybrids were morphologically identified as *N. tessellata*. The majority of hybrids (49) held mitochondrial haplotypes of *N. natrix*, three had haplotypes of *N. tessellata*, and for the remaining snakes the mitochondrial identity is unknown (Table S1). Under $K = 2$, STRUCTURE assigned most samples from Cyprus to the cluster for pure *N. tessellata*, even though the snakes were morphologically unambiguous *N. natrix*. A few other Cypriot samples were genotypically assigned to *N. natrix* or identified as interspecific hybrids. Yet, under $K = 5$, the genotypes of the controversial Cypriot snakes were not lumped together with *N. tessellata* but constituted an extra cluster (Fig. 4 bottom: yellow cluster). Otherwise, *N. tessellata* remained as a distinct cluster, but the former cluster of *N. natrix* was further subdivided, so that there was now one cluster for *N. tessellata*, another one for Cypriot snakes, and three clusters for the remaining *N. natrix*. Under $K = 5$, there was also evidence for interspecific hybridization (Fig. 4; Table S1), even though fewer putative hybrids occurred and all four genotypes of ‘*N. megalcephala*’ were now placed into the cluster of pure *N. natrix*. In a PCA using the same dataset as for the STRUCTURE analyses, *N. tessellata* and *N. natrix* were clearly distinct, although intermediate genotypes occurred. The genotypes of all ‘*N. megalcephala*’ clustered within *N. natrix*. However, most Cypriot snakes were placed completely outside of *N. tessellata* and *N. natrix*; a few Cypriot samples clustered within *N. natrix* (Fig. 5). Population-genetic parameters for all Cypriot samples ($n = 27$) revealed very low genetic variability with evidence for inbreeding compared to other populations harboring the same mitochondrial lineage (i.e., lineage 7; Table S9).

In the face of these conflicting results and the morphological identity of the Cypriot snakes, we used the $K = 5$ scenario as the basis for further calculations. For calculating Q values in HYBRIDLAB to determine interspecific hybrids, the individual Q values of the three clusters for *N. natrix* and of the Cypriot cluster were merged and opposed to the Q values of the cluster for *N. tessellata*. This resulted in thresholds of 93% for pure *N. natrix* and 94% for pure *N. tessellata*.

Genotypic structuring within *Natrix natrix*

Based on these thresholds for pure *Natrix natrix* and *N. tessellata*, a second STRUCTURE run was conducted (Fig. 6A) for which all samples of *N. tessellata* or with genetic impact of *N. tessellata* were excluded. For this dataset ($n = 626$), the ΔK method inferred $K = 2$ as the optimal number of clusters (Fig. S3B). One cluster (green in Fig. 6A) corresponded to populations in the west of our study region. Its geographic distribution matched well that of mtDNA lineage 4 of *N. natrix* (red in the bar for the mi-

1 However, an isolated record of lineage 7 in the western Caucasus region of Turkey (Borçka, Artvin; ZDEU DNA1284) most likely refers to a translocated grass snake or a confused sample. The microsatellite identity of ZDEU DNA1284 matches its mtDNA haplotype, whereas four other snakes from Borçka (ZDEU DNA1283, DNA1287, DNA1288, ZFMK 71145; Table S1) correspond to the expectation for this region and represent the local microsatellite cluster. For one of these snakes the mitochondrial haplotype is known and it belongs to the expected lineage 8.

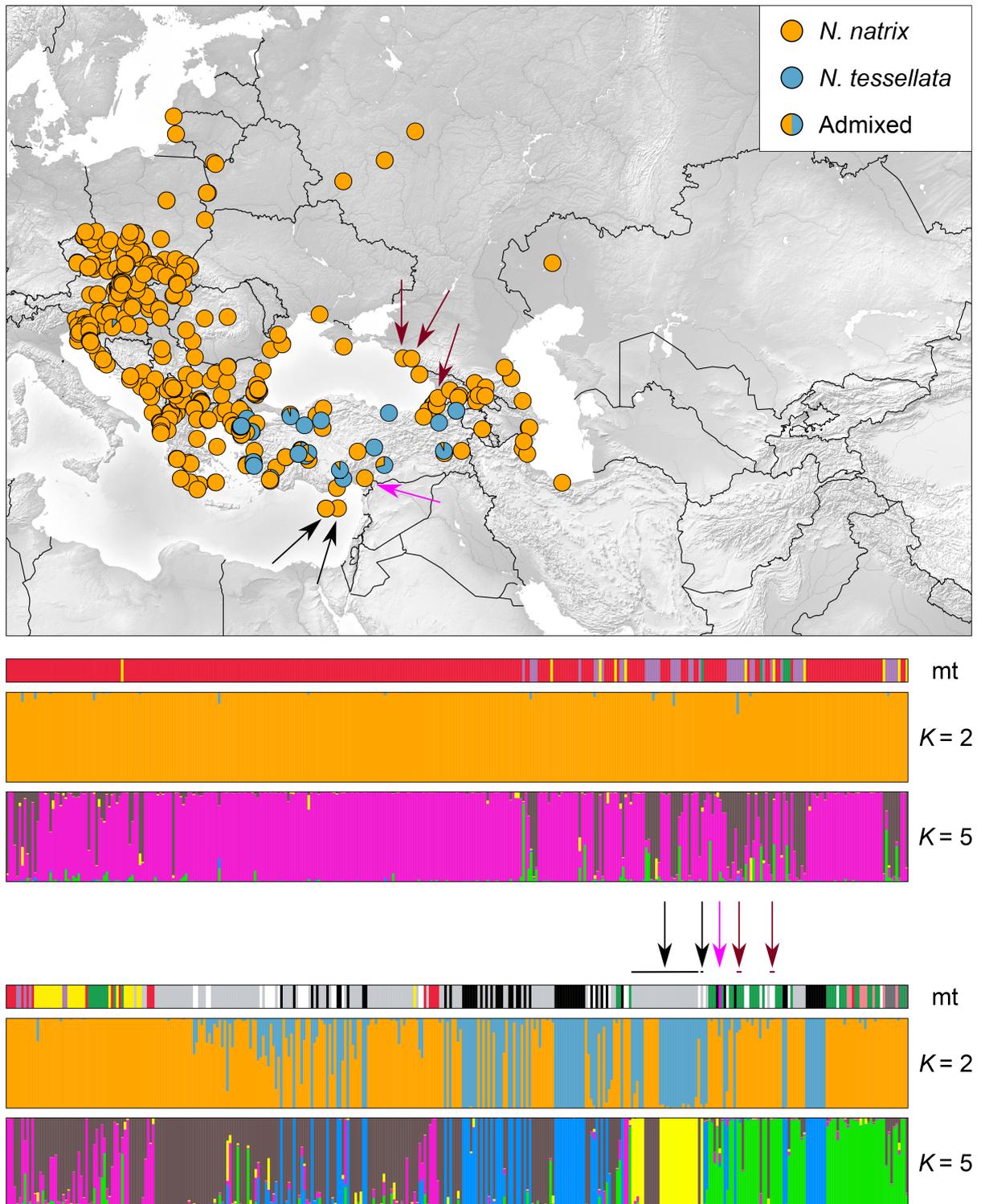


Figure 4. Genotypic structuring of grass snakes and dice snakes in the study region ($n = 705$). In the map, the four clusters representing *Natrix natrix* under $K = 5$ are merged and shown in orange. Admixed ancestries are depicted as pie chart sectors according to Q values as explained in the text. In the bar plots below the map, samples are represented by three bars, indicating their mitochondrial lineage (top) and their inferred STRUCTURE cluster membership under $K = 2$ and $K = 5$ (center and bottom). Mitochondrial lineages of *N. natrix* are color-coded as in Figure 1. However, the two mitochondrial lineages of *N. tessellata* are shown in black to contrast them from *N. natrix*. White indicates missing mtDNA data. Samples are arranged from west to east. In the STRUCTURE diagrams, an individual is represented by a vertical segment that reflects its ancestry. The black arrows highlight samples from Cyprus; the dark red arrows, samples identified as *N. megalcephala*; and the pink arrow indicates the only grass snake with mitochondrial lineage 6 of *N. natrix*.

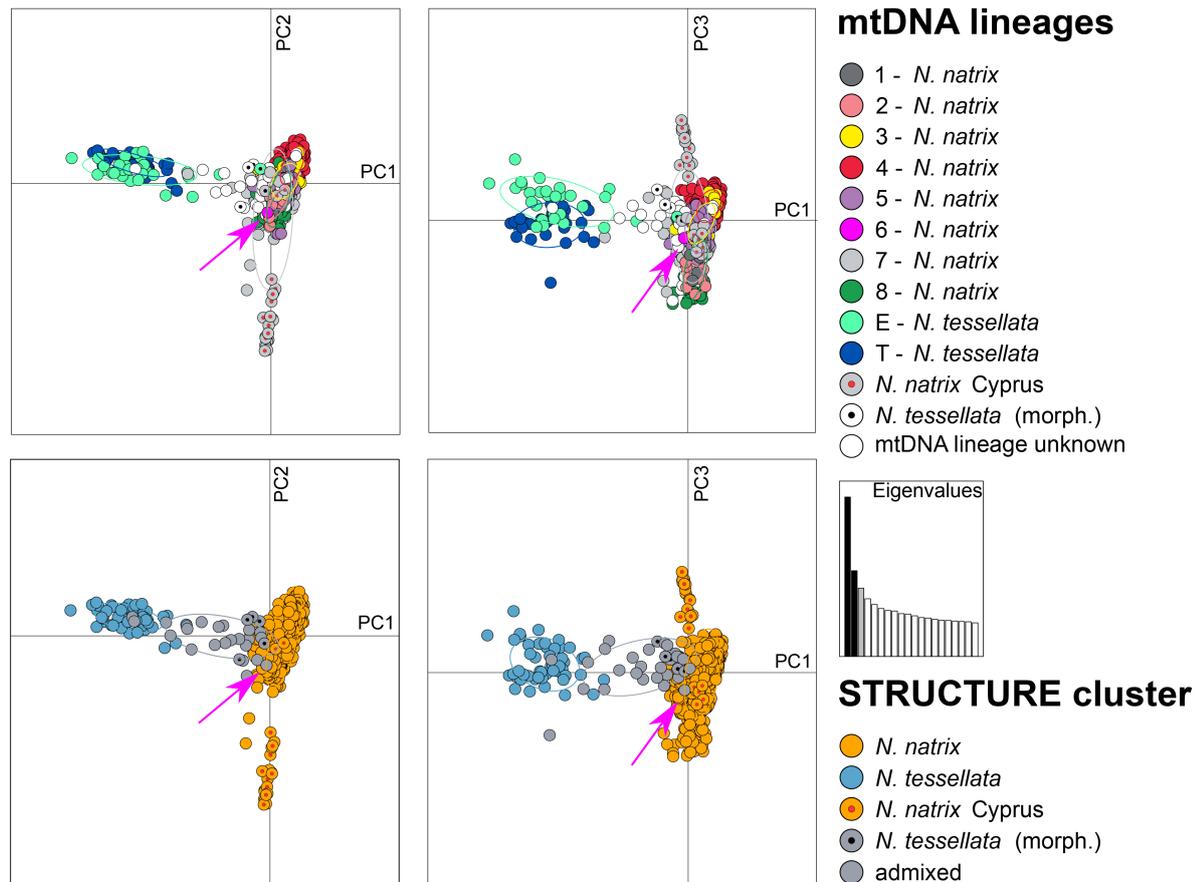


Figure 5. Principal Component Analyses for microsatellite data of all 705 samples (*Natrix natrix*, *N. tessellata*). The same dataset was used as for STRUCTURE analyses. Samples are colored according to their mitochondrial lineage or STRUCTURE cluster membership; admixed individuals in STRUCTURE analyses are shown in grey. Thresholds for admixed samples were derived from HYBRIDLAB simulations and the same as used for STRUCTURE. Oval outlines correspond to 95% confidence intervals. PC1 explains 3.96% of variance, PC2 2.13%, and PC3 1.70%. Some snakes morphologically matching *N. tessellata* are highlighted. The pink arrows denote the only grass snake corresponding to mitochondrial lineage 6 of *N. natrix*.

tochondrial identity in Fig. 6A). The second cluster (pink in Fig. 6A) represented populations in the south and east of our study region and corresponded to several mtDNA lineages distributed there (mtDNA lineages 1, 2, 3, 5, 6, 7, 8; remaining colors in the bar for the mitochondrial identity in Fig. 6A). This cluster also included the four samples from grass snakes morphologically identified as *N. megalcephala* for which microsatellite data are available (ZDEU DNA1287, DNA1288, ZFMK 58029, 60732; Table S1). Admixture between the two clusters occurred mainly in the southern Balkan Peninsula (173 of 205 hybrids) and matched the geographic contact zone of the two clusters. The corresponding PCA showed only weak differentiation, with the western cluster matching mtDNA lineage 4 only slightly differentiated and with massive admixture (Fig. 7A). Yet, many Cypriot samples were again highly distinct.

In the third STRUCTURE analysis, the cluster of *N. natrix* from the south and east, matching several distinct mtDNA lineages, was examined. Samples representing pure *N. natrix* of the western cluster and admixed snakes were removed. For the 159 processed samples, the ΔK method revealed $K = 3$ as the optimal solution (Fig.

S3C). A mainly western cluster (purple in Fig. 6B) largely corresponded to the mitochondrial lineages 3, 5 and 7, whereas an eastern cluster (beige in Fig. 6B) largely matched mitochondrial lineages 1, 2 and 8. A third cluster corresponded to most of the samples from Cyprus (21 of 26 snakes). Ten individuals were identified as admixed. Notably, not all admixed snakes were from the geographic contact zone of the two clusters (except Cyprus cluster). In addition, one sample from Transcaucasia (Azerbaijan) was assigned to the purple cluster (MTD T 3680) and another one from there was admixed (MTD D 48550). The PCA generally also confirmed for this dataset weak differentiation and admixture, but again 21 of the 26 Cypriot samples were clearly distinct (Fig. 7B).

The results of the second and third STRUCTURE analyses largely mirrored the outcome of the first STRUCTURE calculation for $K = 5$ (Fig. 4) in that the four clusters for grass snakes from the first calculation match those also identified in the subset data.

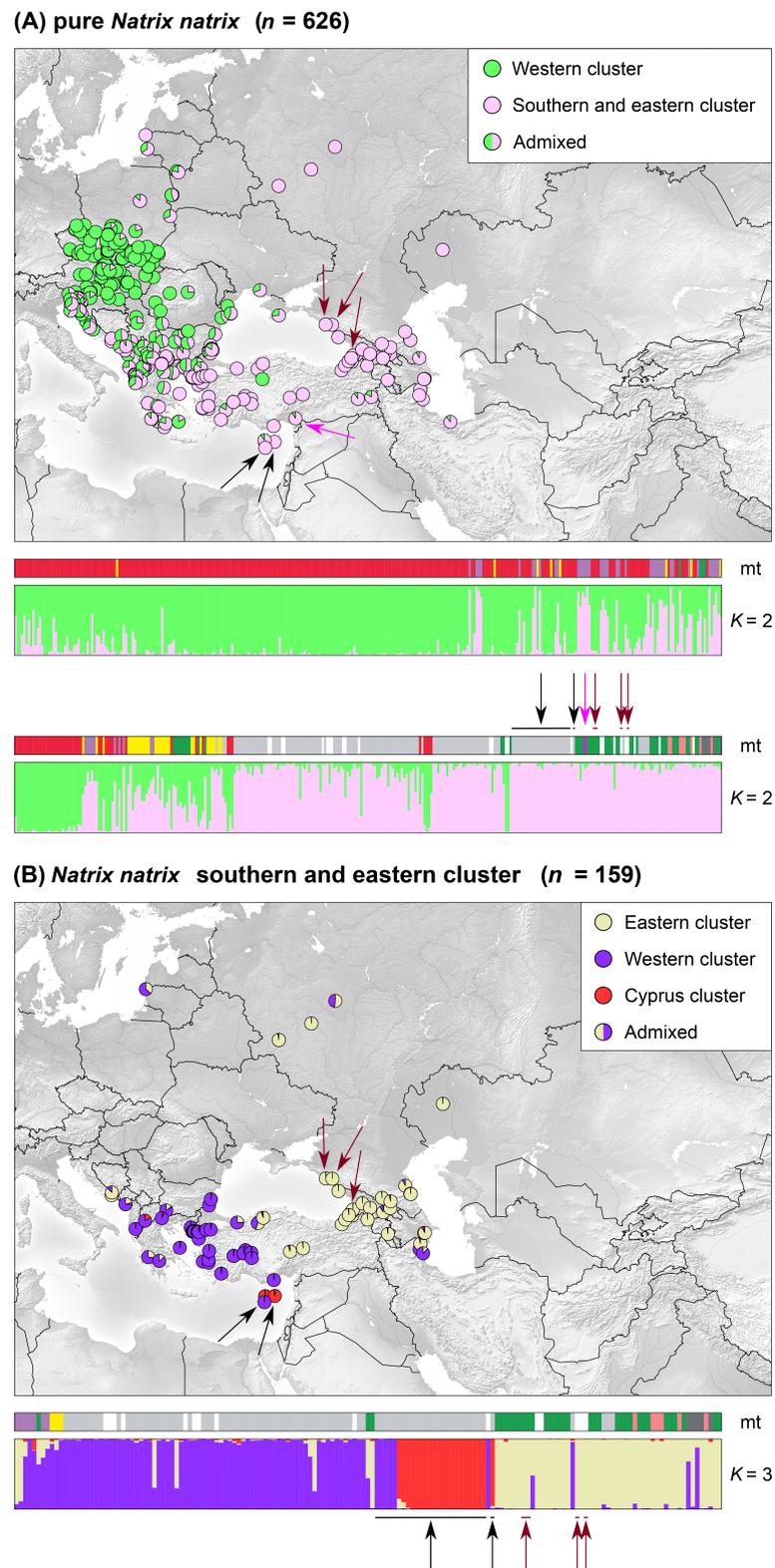


Figure 6. Genotypic structuring of *Natrix natrix* in the study region. Symbol colors indicate STRUCTURE cluster membership; admixed ancestries are depicted as pie chart sectors according to Q values. In the bar plots below the maps, samples are represented by two bars, indicating their mitochondrial lineage (top) and their inferred cluster membership (bottom). Mitochondrial lineages are color-coded as in Figure 1; white bars indicate missing mtDNA data. Samples are arranged from west to east. The black arrows highlight samples from Cyprus; dark red arrows, samples identified as *N. megalcephala*; and the pink arrow indicates the only grass snake with mitochondrial lineage 6 of *N. natrix*. A most likely translocated grass snake from Borçka (Artvin, Turkey) representing the western cluster is not shown in the map (see footnote 1).

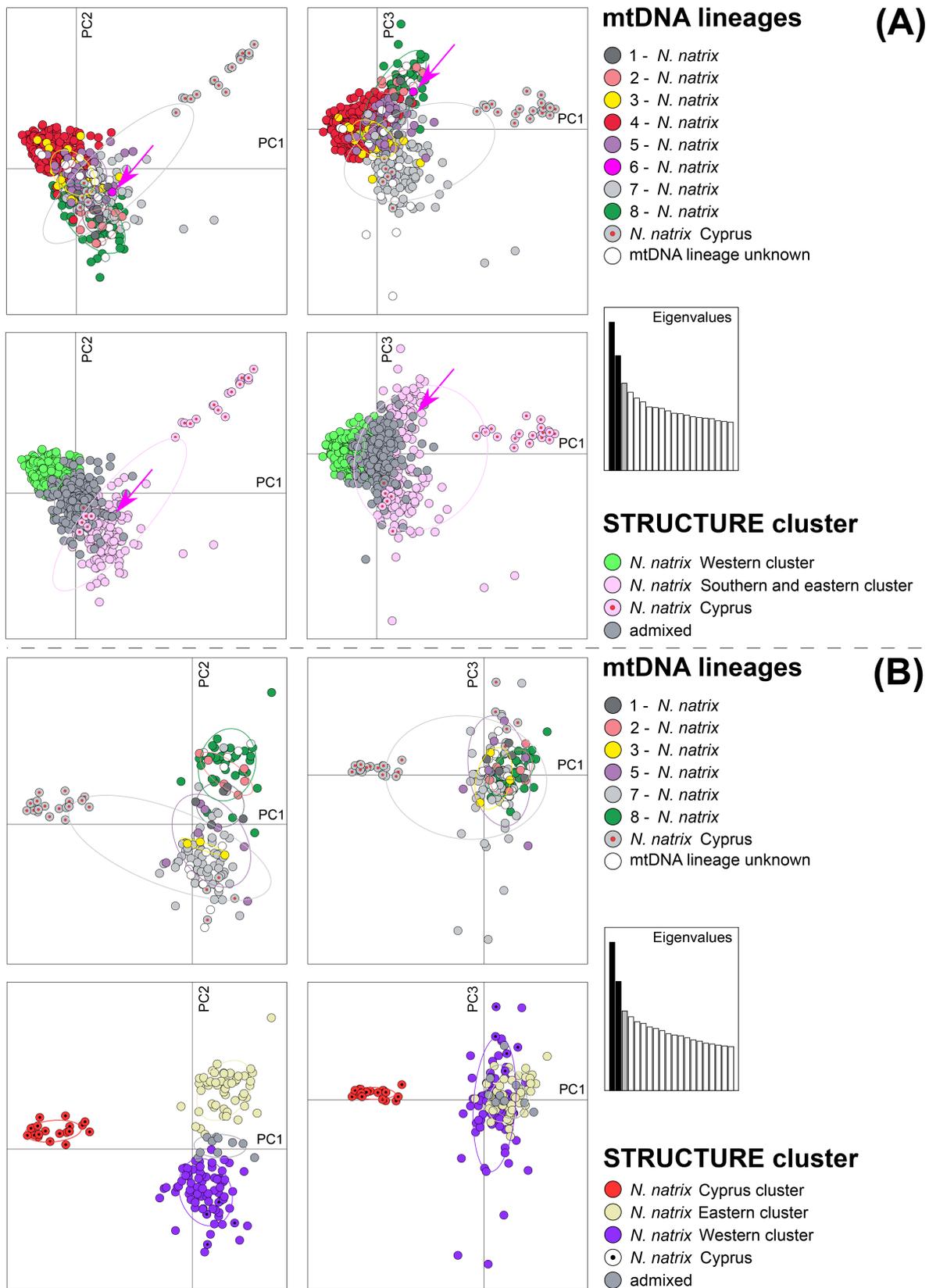


Figure 7. Principal Component Analyses for microsatellite data of (A) pure *Natrix natrix* ($n = 626$) and (B) *N. natrix* of the southern and eastern cluster ($n = 159$). The same datasets were used as for STRUCTURE analyses. Samples are colored according to their mitochondrial lineage or STRUCTURE cluster membership. Thresholds for admixed samples were the same as used for STRUCTURE. Oval outlines correspond to 95% confidence intervals. For (A) PC1 explains 2.82% of variance, PC2 2.18%, and PC3 1.66%; for (B) PC1 explains 4.78% of variance, PC2 3.52%, and PC3 2.57%. The pink arrows highlight a snake corresponding to mitochondrial lineage 6 of *N. natrix*.

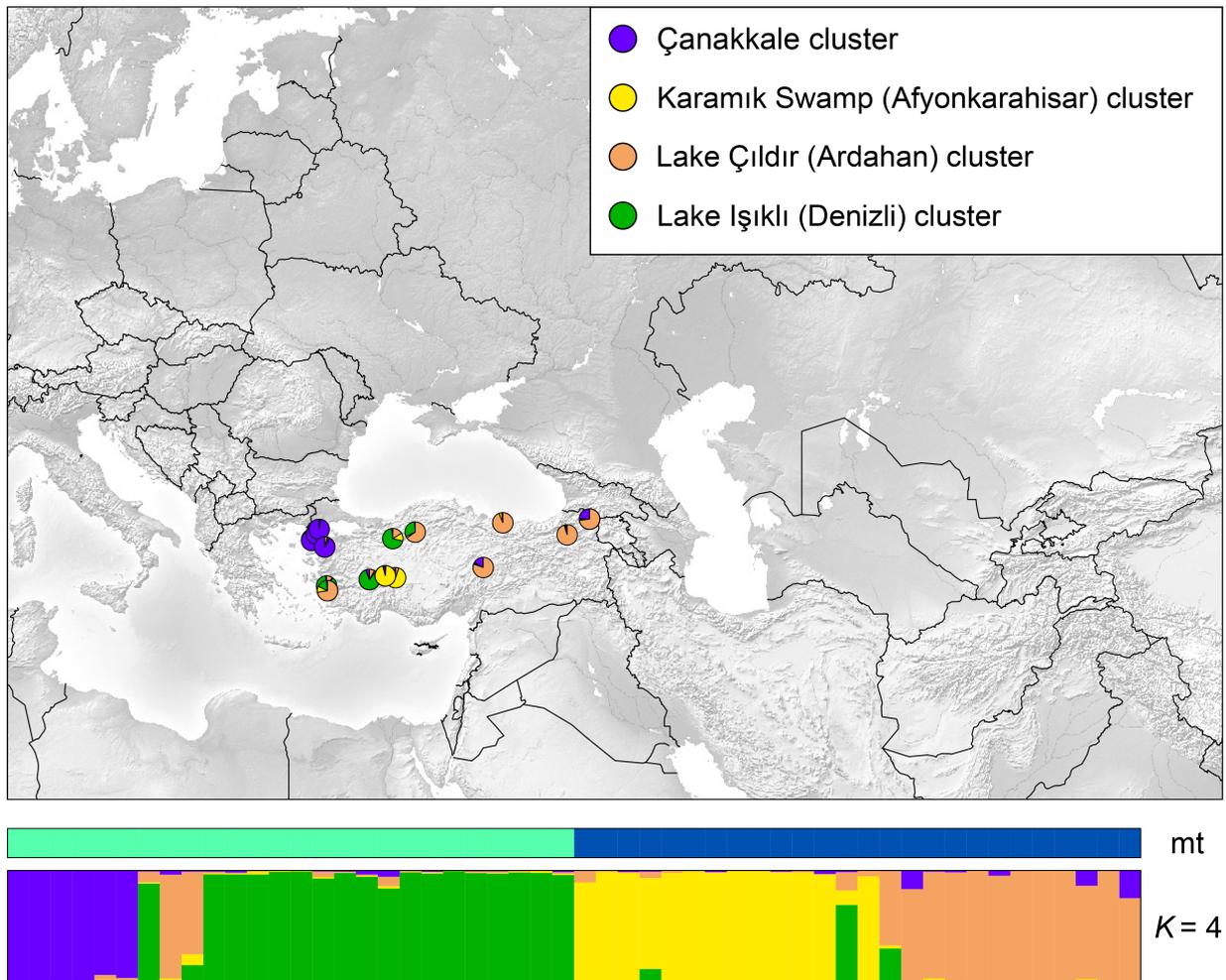


Figure 8. Genotypic structuring of *Natrix tessellata* in the study region ($n = 52$). Symbol colors indicate STRUCTURE cluster membership; admixed ancestries are depicted as pie chart sectors according to Q values. Below the map, the upper bar shows the mitochondrial identity and the lower bar the microsatellite genotype of each individual sample. Mitochondrial lineages color-coded as in Figure 2, with bluish green corresponding to mtDNA lineage E and blue to lineage T. Samples are arranged from west to east.

Genotypic structuring within *Natrix tessellata*

In a fourth STRUCTURE run, only our 52 samples of pure *Natrix tessellata* from Turkey were processed (Fig. 8). For this dataset, the ΔK method inferred $K = 4$ as the optimal number of clusters (Fig. S3D). In the following, we refer to the cluster colors in Figure 8. A purple cluster corresponded to samples from Çanakkale, a green cluster to samples from the Lake Işıklı region (Denizli), a yellow cluster to samples from the Karamık Swamp (Afyonkarahisar), and an orange cluster to most samples from Lake Çıldır (Ardahan) and other sites in eastern Anatolia. The distribution of the orange cluster seems to be disjunct and interrupted by the occurrence of the yellow and green clusters. Admixture seems to be widespread. The purple and the green clusters largely match mtDNA lineage E (bluish green in Fig. 8) and the yellow and orange clusters largely match mtDNA lineage T (blue in Fig. 8), so that two STRUCTURE clusters each correspond to one mito-

chondrial lineage. This differentiation pattern is also confirmed by the PCAs (Fig. 9), which supports both weak differentiation and admixture.

Discussion

The most prominent results of our study are (1) the evidence for hybridization between grass snakes (*Natrix natrix*) and dice snakes (*N. tessellata*), (2) the pronounced divergence of Cypriot grass snakes, (3) the similarity of phylogeographic differentiation patterns of *N. natrix* and *N. tessellata*, and (4) that the eight mitochondrial lineages of *N. natrix* correspond to less microsatellite clusters. These results have implications both for phylogeography in general and for the taxonomy of the two *Natrix* species. In addition, our results confirm that *N. megalcephala* is not a valid species.

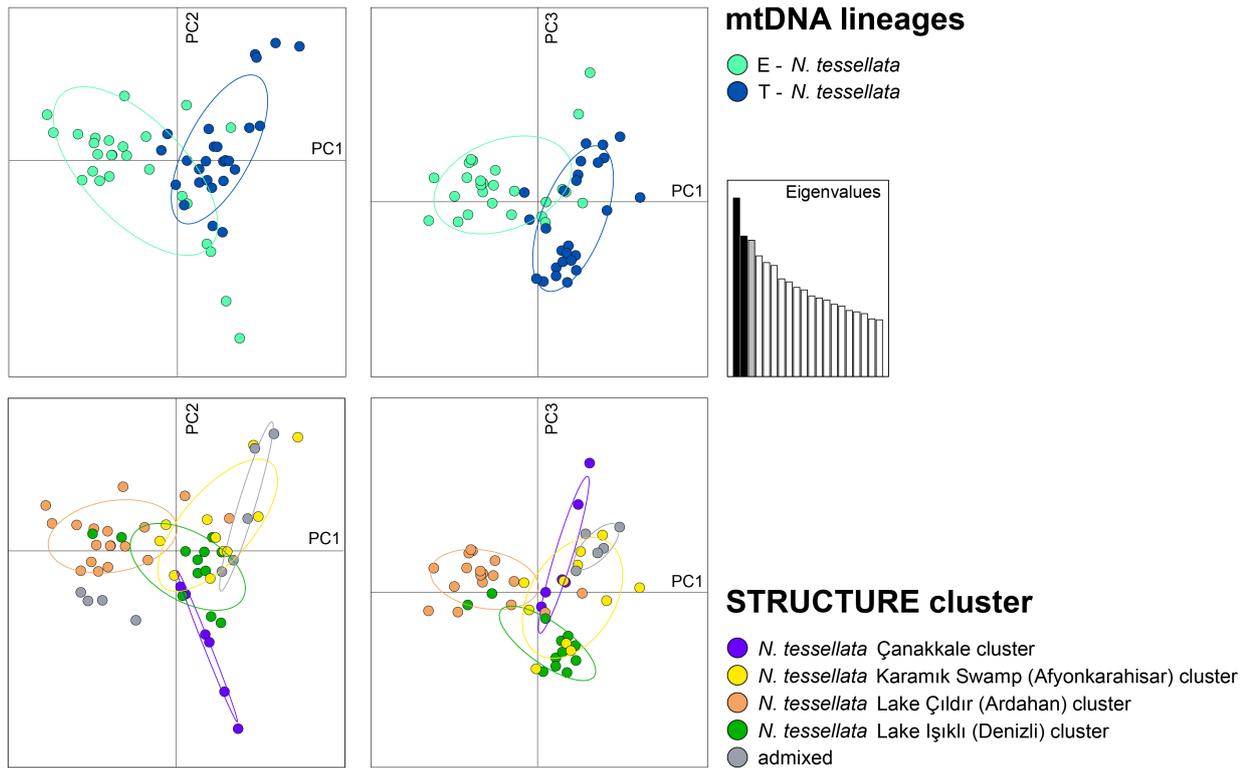


Figure 9. Principal Component Analyses for microsatellite data of pure *Natrix tessellata* ($n = 52$). The same dataset was used as for STRUCTURE analyses. Samples are colored according to their mitochondrial lineage or STRUCTURE cluster membership. Thresholds for admixed samples were the same as used for STRUCTURE. Oval outlines correspond to 95% confidence intervals. PC1 explains 6.65% of variance, PC2 5.23%, and PC3 5.08%.

Interspecific hybridization

Previous investigations focused either on the phylogeography of *N. natrix* (e.g., Kindler et al. 2013, 2017) or *N. tessellata* (Guicking et al. 2009; Guicking and Joger 2011), but the two codistributed species were never examined together. Consequently, these studies were unable to find signals for interspecific hybridization. However, there is a growing body of evidence that interspecific hybridization in animals is more frequent than traditionally thought and may contribute to the enrichment of the gene pools of the involved taxa, in particular via adaptive introgression (Taylor and Larson 2019; Pfennig 2021). Thus, it comes not entirely unexpected that our study revealed evidence for hybridization when using samples of both species, irrespective of whether the microsatellite data were analyzed with STRUCTURE or PCAs (Figs 4 and 5). Why we found interspecific hybridization largely restricted to Turkey remains unclear and calls for further research.

When the putative hybrids from Turkey are compared to the parental species (Table S10), it is obvious that the hybrids share with each parental species many otherwise species-specific alleles. This corroborates their hybrid status, as do specimens that are assigned in the PCAs to the *N. natrix* cluster although they were morphologically identified as *N. tessellata* (Fig. 5; Table S1). Also, three admixed snakes that were morphologically determined

either as *N. natrix* (ZDEU DNA1280) or *N. tessellata* (ZDEU DNA1269, DNA1276) harbored a mitochondrial haplotype of the other species (Table S1), supportive of hybridization. Furthermore, our data indicate that backcrosses occurred (Fig. 4; Table S1), i.e., that the hybrids are obviously not sterile.

Yet, our study also showed the limitations of microsatellites. Using these markers, only recent hybrid and backcross generations can be detected (Sanz et al. 2009). Thus, we cannot trace back older or ancient interspecific gene flow and its extent. Also, a caveat is the possible homoplasy of microsatellites in different species (Pujolar et al. 2014), so that our results need to be scrutinized using other marker systems. Currently, a follow-up investigation using whole genomes is underway. It will also allow the examination of older hybridization events and clarify which parts of the genomes have been exchanged. In any case, our present study exemplifies that phylogeographic investigations disregarding codistributed congeners might miss evidence for interspecific hybridization due to the study design.

Status of Cypriot grass snakes

Regarding Cypriot grass snakes, our study revealed a known weakness of the ΔK method and the STRUCTURE software: The misidentification of $K = 2$ as top level of the

hierarchical structure, even when more clusters should be present (Puechmaile 2016; Janes et al. 2017). This flaw is further exaggerated by uneven sampling (Puechmaile 2016), which is frequently unavoidable. In our case, this situation led to a grave misplacement of most grass snakes from Cyprus. Under $K = 2$, they were lumped together with dice snakes, and only the inspection of $K = 5$ (Fig. 4) and the parallel application of PCAs (Fig. 5) unveiled that this assignment was an artifact. There is only a single allele of the 3TS locus exclusively shared between Cypriot grass snakes and *N. tessellata* (Table S11), so that neither hybridization nor homoplasy can be invoked as reason for the misplacement of Cypriot *N. natrix*. Consequently, the STRUCTURE result under $K = 2$ is clearly erroneous.

Cypriot grass snakes are showing an excess of homozygotes, indicating inbreeding (Table S9). This could have caused their distinctiveness in STRUCTURE analyses under $K = 5$ (Fig. 4) and in PCAs (Fig. 5). Also, mtDNA corroborates that Cypriot grass snakes are genetically impoverished. For *cyt b* and ND4+tRNAs only two haplotypes each were recorded on Cyprus (*cyt b*: gy3 in 18 individuals, gy9 in 6 individuals; ND4+tRNAs: gy6 in 6 individuals, gy11 in 17 individuals), whereas in other sites a greater diversity occurred. For instance on Gökçeada island, being represented by a similar sample size ($n = 17$), four haplotypes each were recorded both for *cyt b* and ND4+tRNAs and in the Karamık Swamp ($n = 10$), four haplotypes for *cyt b* and five for ND4+tRNAs (Table S1). Grass snakes from Gökçeada and the Karamık Swamp share their mitochondrial lineage with Cypriot *N. natrix* (Fig. 3: lineage 7).

Grass snakes are extremely rare on Cyprus and seriously endangered (Blosat 2008; Baier and Wiedl 2010; Zotos et al. 2021). In light of this situation, Cypriot *N. natrix* are best regarded as genetically impoverished, and their distinct microsatellite profile does not imply deep genetic or taxonomic divergence.

The herpetofauna of Cyprus is characterized both by old endemics and recent invaders (Böhme and Wiedl 1994; Poulakakis et al. 2013). The most prominent case of an old endemic species is perhaps the little-known Cyprus racer, *Hierophis cypriensis* (Utiger and Schätti 2004; Kyriazi et al. 2013). Also, the genetic divergences of a toad (*Bufo cypriensis*, Dufresnes et al. 2019), four lizards (*Ablepharus budaki*, *Ophisops elegans*, Poulakakis et al. 2013; *Mediodactylus orientalis*, Kotsakiozi et al. 2018; *Phoenicolacerta troodica*, Tamar et al. 2015) and two snakes (*Telescopus fallax*, Šmíd et al. 2019; *Xerophyllops vermicularis*, Kornilios 2017) qualify them as old Cypriot endemics, whereas the shallow genetic differences of some other amphibians (*Hyla savignyi*, *Pelophylax* sp., Poulakakis et al. 2013) and reptiles (*Mauremys rivulata*, Vamberger et al. 2017; *Acanthodactylus schreiberi*, Poulakakis et al. 2013; *Chalcides ocellatus*, Bisbal-Chinesta et al. 2020; *Hemidactylus turcicus*, Rato et al. 2011; *Rhynchocalamus melanocephalus*, Tamar et al. 2020) suggest that these taxa are recent invaders, even though this view has been challenged for the water frogs (*Pelophylax* sp., Plötner et al. 2012; but see Speybroeck et al. 2020). The immigration of the recent invaders seems to

be frequently human-mediated (Böhme and Wiedl 1994; Poulakakis et al. 2013).

The local grass snakes have been treated as another candidate for an old Cypriot endemic (Böhme and Wiedl 1994). However, Cypriot grass snakes share their mitochondrial lineage with *N. natrix* from the southeastern Balkan Peninsula and western and southern Anatolia (Fig. 3: lineage 7). When the individual haplotypes are inspected, it becomes obvious that three of the four haplotypes found on Cyprus also occur in Bulgaria, Greece, and Turkey (*cyt b*: gy9; ND4+tRNAs: gy6, gy11; Table S1). If the grass snakes would represent an old Cypriot endemic, we would expect a deep mitochondrial separation from the mainland populations. This is not the case and the present evidence instead supports recent immigration and subsequent genetic impoverishment.

Comparative phylogeographies of *Natrix natrix* and *N. tessellata*

In general, the mitochondrial phylogeographies of *N. natrix* and *N. tessellata* show some similarities (Fig. 3), both with respect to the distribution of some mitochondrial lineages and the location of contact zones, suggestive of a shared biogeographic history. For both species, the Balkans, Anatolia or Western Asia in general, and the Caucasus have undoubtedly served as glacial refugia, where genetic lineages survived that diverged prior to the Pleistocene. For *N. natrix*, a fossil-calibrated molecular clock estimated that six lineages separated 3.4–7.2 million years ago, long before the Pleistocene glaciations, and only two of these lineages diverged then further during the early Pleistocene (1.5 and 1.7 million years ago; Kindler et al. 2018a). Yet, the current distribution pattern undoubtedly has been massively shaped by repeated glacial range restrictions and interglacial range expansions, and the wide distribution of one mitochondrial lineage in the northeast of the distribution ranges of the two species (Fig. 3: lineage 8 in *N. natrix*; lineage A in *N. tessellata*) clearly results from Holocene range expansions. Thus, the general phylogeographic patterns in both species match the well-known paradigm of southern genetic richness to northern purity (Hewitt 2000), reflecting the survival of distinct lineages in former southern glacial refugia and the Holocene colonization of more northerly regions by only one or a few of these lineages.

In contrast to *N. natrix*, *N. tessellata* shows additional lineages in regions where grass snakes do not occur, i.e., on the island of Crete, in the southern Levant and Egypt or in southern Central Asia. It remains a challenge to study the disjunct populations of *N. natrix* from Siberia and the Lake Baikal region to find out how they match the general pattern.

As in many other taxa (Joger et al. 2007; Schmitt 2020), multiple lineages, which most likely can be traced back to distinct glacial microrefugia, occur in the Balkan Peninsula and in the Caucasus region, both in *N. natrix* and *N. tessellata*. In these two regions, and further north in the Balkan Peninsula and in Anatolia, also lie

secondary contact zones that can be understood as having formed during Holocene range expansions.

In dice snakes, we are lacking nuclear genomic evidence for most lineages. Our sample is restricted to Turkey. Yet, it shows that the two mitochondrial lineages occurring there (Fig. 3: E, T) match two microsatellite clusters each, i.e., that mitochondrial divergence parallels nuclear genomic differentiation. Further, there is clear evidence of admixture (Figs 8 and 9), suggestive of conspecificity. However, the deep mitochondrial divergences among all mitochondrial lineages of *N. tessellata* (Tables S7, S8) resemble those observed between *N. astreptophora*, *N. helvetica*, and *N. natrix* (Kindler et al. 2017, 2018a), i.e., three species with much restricted gene flow. This implies that dice snakes could also represent more than only one species. For a better understanding of their taxonomy, microsatellite data or other nuclear genomic information is needed beyond Turkey.

For grass snakes, range-wide microsatellite data are available (Figs 6 and 7). As in *N. tessellata*, there is evidence for admixture in the contact zones of distinct microsatellite clusters. Contrary to *N. tessellata*, in *N. natrix* the number of mitochondrial lineages, however, exceeds the number of microsatellite clusters (Fig. 3). Eight mitochondrial lineages correspond to only four microsatellite clusters, and if the Cyprus cluster is disregarded, only to three (Figs 3, 6 and 7). The uppermost hierarchical level of nuclear genomic divergence is represented by one microsatellite cluster from the west of our study region that matches mtDNA lineage 4. Another southern and eastern microsatellite cluster corresponds to several distinct mtDNA lineages (Figs 6A and 7A). This latter southern and eastern microsatellite cluster is further structured in three clusters when analyzed alone: Two western clusters (Cyprus versus southern Balkans plus western Turkey) largely match mtDNA lineages 3, 5, and 7, and a third eastern cluster corresponds to three other mtDNA lineages (1, 2, 8). This third cluster also includes all four samples morphologically identified as *N. megalcephala* (Fig. 6B). Further sampling from the southern Balkans and the eastern distribution range is needed to disentangle this situation in more detail. We cannot rule out that more samples will unveil more structure than our present data set. However, the PCA for the grass snakes from the south and east (Fig. 7B) shows quite compact scatter plots that rather argue for genetic homogeneity of each of the three clusters (with the inbred Cypriot population being most distinct).

How can this be explained? A similar situation is known for the closely related barred grass snake (*N. helvetica*). Grass snakes are highly mobile animals, with males having larger home ranges than females (Madsen 1984; Wisler et al. 2008; Reading and Jofré 2009; Meister et al. 2012). In Italy, five in part deeply divergent mtDNA lineages of *N. helvetica* correspond to only a single nuclear genomic cluster, due to largely male-mediated gene flow and female philopatry (Schultze et al. 2020). We hypothesize that the situation in our study region parallels that in Italy, i.e., that the distinct mtDNA lineages corresponding to one and the same microsatellite cluster are vestiges of former genetic divergence that is increasingly

eradicated due to recent gene flow, which commenced when the snakes dispersed out of distinct glacial microrefugia in the Holocene. So to speak, the Holocene range expansions transformed the glacial ‘hotspots’ into ‘melting pots,’ in which only ‘mitochondrial ghost lineages’ bear witness of the formerly distinct evolutionary lineages. Such a scenario explains both the situation on the southern Balkan Peninsula and in the Caucasus region. There, distinct mitochondrial lineages occur in close geographic proximity but correspond to only one microsatellite cluster each (Figs 3, 6 and 7: Balkans – mtDNA lineages 3, 5, 7; Caucasus – mtDNA lineages 1, 2, 8).

For the Balkan Peninsula and southern Central Europe, this also suggests that a ‘rolling expansion and hybridization wave’ established when grass snakes with mtDNA lineage 4 (*N. n. vulgaris*) spread northwestward. While *N. n. vulgaris* successively invaded more northerly regions and ‘genetically swamped’ the resident grass snakes characterized by mtDNA lineage 3 (*N. n. natrix*; see Asztalos et al. 2021) there, the populations of *N. n. vulgaris* in their former southern refuge became increasingly admixed when other lineages, having independent glacial refuges in the southern Balkan Peninsula, expanded their ranges with Holocene warming. This led to the current, at first glance counterintuitive, pattern that no pure populations of *N. n. vulgaris* exist anymore in their former glacial refuge, in contrast to their Holocene expansion area, where relatively pure populations occur in southern Central Europe and the adjacent northern Balkans.

Among the mtDNA lineages that had glacial refugia in the southern Balkan Peninsula is also mtDNA lineage 3, otherwise characteristic for the nominotypical subspecies *N. n. natrix* (see Fritz and Schmidler 2020). However, the disjunct Balkan haplotypes slightly differ from the northern ones found within the range of *N. n. natrix*, suggesting that lineage 3 represents on the southern Balkan Peninsula an old genetic relict of a formerly wider distribution during an earlier interglacial. In contrast, the distinct Central European and Scandinavian haplotypes of lineage 3 provide evidence for a northern glacial refuge in Central Europe (Kindler et al. 2018b).

Taxonomic implications

Our results corroborate that the species *Natrix megalcephala* is invalid (Kindler et al. 2013; Fritz and Schmidler 2020) because it is not genetically differentiated from populations of *N. natrix* in the same geographic region. With respect to *N. tessellata*, the congruence of nuclear genomic and mitochondrial divergence in our study region suggests that the earlier described deeply divergent mitochondrial lineages (Guicking et al. 2009; Guicking and Joger 2011) reflect taxonomic differentiation. Evidence for gene flow between the two lineages in our study region supports their conspecificity, if large-scale gene flow and reproductive isolation are used as criteria for species delineation (cf. Fritz and Schmidler 2020; Speybroeck et al. 2020). Also for grass snakes the admixture revealed in our analyses supports conspecificity.

For dice snakes additional genetic and morphological investigations are needed for a comprehensive taxonomic revision. For *N. natrix* the present evidence and the nomenclatural foundation laid by Fritz and Schmidtler (2020) allow assigning names to the conspecific clusters and refining the currently accepted subspecies classification. In doing so, we follow Kindler and Fritz (2018) and understand subspecies as evolutionarily significant units that differ in their nuclear genomic and mitochondrial identity and that are, in contrast to species, fully capable of extensive gene flow that can even lead to complete genetic and taxonomic amalgamation. If mitochondrial ghost lineages (i.e., higher mitochondrial diversity), mitochondrial introgression or mitochondrial capture (i.e., lower mitochondrial diversity and mitonuclear discordance) occur, nuclear genomic differentiation is decisive for subspecies delimitation.²

Within this framework (Fig. 10), the microsatellite cluster from the west of our study region (characterized by mtDNA lineage 4) is to be identified with the subspecies *Natrix natrix vulgaris* Laurenti, 1768 (Fritz and Schmidtler 2020). The nomenclatural situation for the south and east of our study region is less straightforward and complicated by extensive past and current hybridization, leading to the survival of many mitochondrial ghost lineages that conflict with nuclear genomic differentiation. In this region, traditionally many subspecies were recognized (Kabisch 1999; Kreiner 2007; Geniez 2015), among them also *Natrix natrix cyprica* (Hecht, 1930). However, acknowledging that Cypriot grass snakes are distinct on the nuclear genomic level only due to genetic impoverishment, we propose that Cypriot populations should be taxonomically lumped together with grass snakes from the southern Balkans and western Anatolia characterized by mtDNA lineage 7. Using microsatellites, these populations cluster together with grass snakes from the southern Balkans also harboring mtDNA lineages 3 and 5. The oldest available name for these populations is *Natrix natrix moreotica* (Bedriaga, 1882) with a type locality on the northern Peloponnese (Greece; see Fritz and Schmidtler 2020). This renders *Natrix natrix cyprica* (Hecht, 1930) from Cyprus a junior synonym of *N. n. moreotica*.

Grass snakes of the eastern microsatellite cluster, characterized by mtDNA lineages 1, 2, and 8, are currently assigned to two subspecies (Fritz and Schmidtler 2020). *Natrix natrix scutata* (Pallas, 1771), with a type locality of Atyrau (Kazakhstan), is identified with the northern populations, while the name *Natrix natrix persa* (Pallas, 1814), type locality Gilan and Mazandaran (Iran), is assigned to the southern populations. Our present results suggest that these two taxa should be synonymized and that the older name, *N. n. scutata*, should be used for all populations.

Furthermore, our results are inconclusive with respect to another subspecies that was recognized as val-

id by Fritz and Schmidtler (2020), namely *Natrix natrix syriaca* (Hecht, 1930), with a type locality of Zincirli in southeastern Turkey, near the Gulf of İskenderun. Fritz and Schmidtler (2020) tentatively identified this subspecies with grass snakes harboring mtDNA lineage 6. We could study only microsatellite data of a single individual of this lineage and this situation prevents any taxonomic conclusions, even though this sample did not seem to be distinct from other material (Figs 4–7).

Conclusions

Our study showcases that phylogeographic investigations should include samples of codistributed congeners because otherwise evidence for interspecific hybridization will be missed. Genomic approaches are expected to be more informative than investigations based on microsatellite loci alone. In our study system, we identified hybridization between *Natrix natrix* and *N. tessellata*, two species that were previously thought to hybridize only rarely (Mebert et al. 2011). Our study also exemplifies that the study of codistributed taxa allows the direct comparison of phylogeographic differentiation patterns and more general inferences about biogeography. Furthermore, our findings support that the software STRUCTURE is prone to errors regarding the number of clusters when no additional evidence, independent from population-genetic assumptions, is considered. With respect to *N. tessellata* our results suggest that the previously identified mitochondrial lineages reflect taxonomic divergence, as in *N. natrix*. Further research is needed in *N. tessellata* to examine nuclear genomic variation beyond Turkey. The deep mitochondrial lineages in this species (Guicking et al. 2009; Guicking and Joger 2011) and the nuclear genomic fingerprints presented in Guicking et al. (2009) suggest that more than only a single species could be involved, although this is not confirmed for the two mitochondrial lineages studied herein. For *N. natrix*, we propose a refined intraspecific classification and recognize three subspecies from our study region, bringing the total number of subspecies to four (Fig. 10): *Natrix natrix vulgaris* Laurenti, 1768 from southeastern Central Europe and the northern Balkans, *Natrix natrix moreotica* (Bedriaga, 1882) from the southern Balkans, western Anatolia and Cyprus, and *Natrix natrix scutata* (Pallas, 1771) from eastern Anatolia, the Caucasus region, Iran, and a wide area from eastern Poland and Finland to Kazakhstan and the Lake Baikal region in Russia and adjacent Mongolia. All of these three subspecies comprise individuals with and without back stripes, with an increasing tendency to the presence of back stripes and other coloration variations in the south of the range. The status of *Natrix natrix syriaca* (Hecht, 1930) from the Gulf of İskenderun, Turkey, cannot be resolved due to insufficient material. Beyond our study region, *Natrix natrix natrix* (Linnaeus, 1758) inhabits the northwestern distribution range of the species (Central Europe and Scandinavia).

2 For instance, mitochondrial ghost lineages were described for *N. helvetica sicula* in mainland Italy and Sicily (Schultze et al. 2020). Many populations of *N. n. vulgaris* in southern Germany captured mtDNA haplotypes of *N. n. natrix* (Asztalos et al. 2021a).

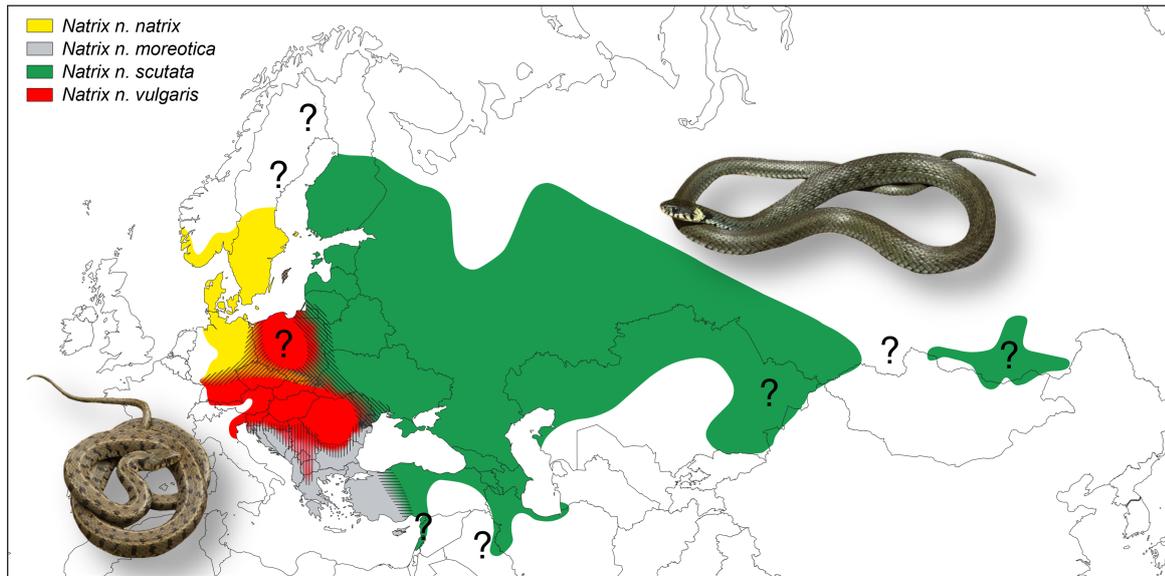


Figure 10. Approximate distribution of the subspecies of *Natrix natrix* as outlined in the present study, Fritz and Schmidler (2020), and Asztalos et al. (2021a). Hatching indicates zones of secondary contact and admixture. *Natrix natrix natrix* (Linnaeus, 1758) largely matches mtDNA lineage 3 [non-Balkan haplotypes, see Discussion]; *N. n. vulgaris* Laurenti, 1768, mtDNA lineage 4; *N. n. moreotica* (Bedriaga, 1882), mtDNA lineages 3 [Balkan haplotypes, see Discussion], 5, and 7; and *N. n. scutata* (Pallas, 1771), mtDNA lineages 1, 2, and 8. Note that conflicting mitochondrial lineages occur in areas of mitonuclear discordance (e.g., many populations of *N. n. vulgaris* in southern Germany have haplotypes of mtDNA lineage 3). Question marks denote unclear taxonomic identity and/or unclear range borders. Map is based on records from Bannikov et al. (1977), Fog et al. (1997), Milto (2003), Grilitsch and Werner (2009), Kuranova et al. (2010), Afrasiab et al. (2011), Litvinchuk et al. (2013), Sindaco et al. (2013), GBIF (www.gbif.org), and iNaturalist (<https://www.inaturalist.org>), combined with our genetically verified records, following the approach described in Asztalos et al. (2020). Insets: *Natrix natrix moreotica* (bottom left; Peloponnese, Greece) and *N. n. scutata* (top right; Białowieża, easternmost Poland). Photos: Henrik Bringsøe.

Acknowledgements

The authors thank the Scientific and Technological Research Council of Turkey TÜBİTAK (project number 116Z359) for financial support. Marika Asztalos' work was funded by Senckenberg. Daniel Jablonski 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. All animal handling within the scope of the TÜBİTAK project was in accordance with wildlife husbandry standards approved by Decision 2016–18 of the Laboratory Animals Ethical Committee at Ege University, Turkey. Süleyman İlhan and Melodi Yenmiş assisted with laboratory work. Frank Glaw and Ben Wielstra donated some samples. Henrik Bringsøe allowed the use of two photos for a figure. Frank Glaw and an anonymous reviewer made helpful suggestions that improved an earlier version of this study. Flora Ihlow and Markward Herbert Fischer produced one map for us.

References

- Afrasiab SR, Al-Ganabi MI, Al-Fartosi K (2011) Snake species new or rare to the herpetofauna of Iraq. *Herpetozoa* 24: 179–181.
- Ahnelt H, Romanova T, Klinge A, Böhme W, Fritz U, Asztalos M (2021) The common grass snake (*Natrix natrix*) on Sylt: Human-mediated colonization of a North Sea island. *Salamandra* 57: 285–290.
- Asztalos M, Glaw F, Franzen M, Kindler C, Fritz U (2021a) Transalpine dispersal: Italian barred grass snakes in southernmost Bavaria – This far but no further! *Journal of Zoological Systematics and Evolutionary Research* 59: 1136–1148. <https://doi.org/10.1111/jzs.12471>
- Asztalos M, Schultze N, Ihlow F, Geniez P, Berroneau M, Delmas C, Guillier G, Legentilhomme J, Kindler C, Fritz U (2020) How often do they do it? An in-depth analysis of the hybrid zone of two grass snake species (*Natrix astreptophora*, *N. helvetica*). *Biological Journal of the Linnean Society* 131: 756–773. <https://doi.org/10.1093/biolinnean/blaa152>
- Asztalos M, Wielstra B, Struijk RPJH, Ayaz D, Fritz U (2021b) Aliens in the Netherlands: Local genetic pollution of barred grass snakes (Squamata: Serpentes: Natricidae). *Salamandra* 57: 174–179.
- Baier F, Wiedl H-J (2010) The re-evaluated conservation status of the mountain populations of the highly endangered Cyprus grass snake, *Natrix natrix cypriaca* (Hecht, 1930), with miscellaneous natural history notes. *Salamandra* 46: 16–23.
- Банников АГ, Даревский ИС, Ищенко ВГ, Рустамов АК, Щербак НН [Bannikov AG, Darevsky IS, Ishchenko VG, Rustamov AK, Shcherbak NN] (1977) Определитель земноводных и пресмыкающихся фауны СССР. Просвещение, Москва, 415 pp.
- Bisbal-Chinesta JF, Tamar K, Gálvez Á, Albero L, Vicent-Castelló P, Martín-Burgos L, Alonso M, Sánchez R, Ortega C, Gómez A, Candel D, Cervera M, Carranza S, Blain H-A (2020) Trade and stowaways: Molecular evidence for human-mediated translocation of eastern skinks into the western Mediterranean. *Amphibia-Reptilia* 41: 49–62. <https://doi.org/10.1163/15685381-20191249>
- Blosat B (2008) Population status, threats and protection of the grass snake, *Natrix natrix cypriaca* (Hecht, 1930) on Cyprus. *Mertensiella* 17: 246–271.

- Böhme W, Wiedl H (1994) Status and zoogeography of the herpetofauna of Cyprus, with taxonomic and natural history notes on selected species (genera *Rana*, *Coluber*, *Natrix*, *Vipera*). *Zoology in the Middle East* 10: 31–52.
- Clement M, Posada D, Crandall KA (2000) tcs: A computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657–1659. <https://doi.org/10.1046/j.1365-294x.2000.01020.x>
- Dufresnes C, Mazepa G, Jablonski D, Caliani Oliveira R, Wenseleers T, Shabanov DA, Auer M, Ernst R, Koch C, Ramirez-Chaves HE, Mulder KP, Simonov E, Tiutenko A, Kryvokhyzha D, Wenkes PL, Zinenko OI, Korshunov OV, Al-Johany AM, Peregontsev EA, Masroor R, Betto-Colliard C, Denoël M, Borkin LJ, Skorinov DV, Pasyukova RA, Mazanaeva LF, Rosanov JM, Dubey S, Litvinchuk SN (2019) Fifteen shades of green: The evolution of *Bufo* toads revisited. *Molecular Phylogenetics and Evolution* 141: 106615. <https://doi.org/10.1016/j.ympev.2019.106615>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Lischer HEL (2010) ARLEQUIN suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587. <https://doi.org/10.1093/genetics/164.4.1567>
- Fog K, Schmedes A, Rosenørn de Lasson D (1997) Nordens padder og krybdyr. G.E.C. Gad, Copenhagen, 365 pp.
- Francis RM (2017) POPHELPER: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* 17: 27–32. <https://doi.org/10.1111/1755-0998.12509>
- Fritz U, Schmidler JF (2020) The Fifth Labour of Heracles: Cleaning the Linnean stable of names for grass snakes (*Natrix astreptophora*, *N. helvetica*, *N. natrix* sensu stricto). *Vertebrate Zoology* 70: 621–665. <https://doi.org/10.26049/VZ70-4-2020-07>
- Geniez P (2015) Serpents d'Europe, d'Afrique du Nord et du Moyen-Orient. Delachaux et Niestlé, Paris, 380 pp.
- Goudet J (1995) FSTAT (version 1.2): A computer program to calculate *F*-statistics. *Journal of Heredity* 86: 485–486. <https://doi.org/10.1093/oxfordjournals.jhered.a111627>
- Grillitsch H, Werner YL (2009) The southern limit of *Natrix natrix* in the Levant – A detective story (Squamata: Serpentes: Colubridae). *Herpetozoa* 22: 65–74.
- Gruschwitz M, Lenz S, Mebert K, Laňka V (1999) *Natrix tessellata* (Laurenti, 1768) – Würfelnatter. In: Böhme W (Ed.) *Handbuch der Reptilien und Amphibien Europas*. Band 3/IIA: Schlangen II. Aula-Verlag, Wiebelsheim, 581–644.
- Guicking D, Joger U (2011) Molecular phylogeography of the dice snake. *Mertensiella* 18: 1–10.
- Guicking D, Joger U, Wink M (2008) Molecular phylogeography of the viperine snake *Natrix maura* (Serpentes: Colubridae): Evidence for strong intraspecific differentiation. *Organisms, Diversity & Evolution* 8: 130–145. <https://doi.org/10.1016/j.ode.2007.05.001>
- 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. *Organisms, Diversity & Evolution* 9: 201–214. <https://doi.org/10.1016/j.ode.2009.03.001>
- Guicking D, Lawson R, Joger U, Wink M (2006) Evolution and phylogeny of the genus *Natrix* (Serpentes: Colubridae). *Biological Journal of the Linnean Society* 87: 127–143. <https://doi.org/10.1111/j.1095-8312.2006.00561.x>
- Hall TA (1999) BIOEDIT: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913. <https://doi.org/10.1038/35016000>
- Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorell JC, Cullingham CI, Andrew RL (2017) The *K* = 2 conundrum. *Molecular Ecology* 26: 3594–3602. <https://doi.org/10.1111/mec.14187>
- Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution* 15: 250–255. [https://doi.org/10.1016/s0169-5347\(00\)01873-5](https://doi.org/10.1016/s0169-5347(00)01873-5)
- Joger U, Fritz U, Guicking D, Kalyabina-Hauf S, Nagy ZT, Wink M (2007) Phylogeography of western Palaearctic reptiles – Spatial and temporal speciation patterns. *Zoologischer Anzeiger* 246: 293–313. <https://doi.org/10.1016/j.jcz.2007.09.002>
- Jombart T (2008) ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Kabisch K (1999) *Natrix natrix* (Linnaeus, 1758) – Ringelnatter. In: Böhme W (Ed.) *Handbuch der Reptilien und Amphibien Europas*. Band 3/IIA: Schlangen II. Aula-Verlag, Wiebelsheim, 513–580.
- Kindler C, Böhme W, Corti C, Gvoždík V, Jablonski D, Jandzik D, Metallinou M, Široký P, Fritz U (2013) Mitochondrial phylogeography, contact zones and taxonomy of grass snakes (*Natrix natrix*, *N. megaloccephala*). *Zoologica Scripta* 42: 458–472. <https://doi.org/10.1111/zsc.12018>
- Kindler C, Bringsøe H, Fritz U (2014) Phylogeography of grass snakes (*Natrix natrix*) all around the Baltic Sea: Implications for the Holocene colonization of Fennoscandia. *Amphibia-Reptilia* 35: 413–424. <https://doi.org/10.1163/15685381-00002962>
- Kindler C, Chèvre M, Ursenbacher S, Böhme W, Hille A, Jablonski D, Vamberger M, Fritz U (2017) Hybridization patterns in two contact zones of grass snakes reveal a new Central European snake species. *Scientific Reports* 7: 7378. <https://doi.org/10.1038/s41598-017-07847-9>
- Kindler C, de Pous P, Carranza S, Beddek M, Geniez P, Fritz U (2018a) Phylogeography of the Ibero-Maghrebian red-eyed grass snake (*Natrix astreptophora*). *Organisms Diversity & Evolution* 18: 143–150. <https://doi.org/10.1007/s13127-017-0354-2>
- Kindler C, Fritz U (2018) Phylogeography and taxonomy of the barred grass snake (*Natrix helvetica*), with a discussion of the subspecies category in zoology. *Vertebrate Zoology* 68: 269–281.
- Kindler C, Graciá E, Fritz U (2018b) Extra-Mediterranean glacial refuges in barred and common grass snakes (*Natrix helvetica*, *N. natrix*). *Scientific Reports* 8: 1821. <https://doi.org/10.1038/s41598-018-20218-2>
- Kornilios P (2017) Polytomies, signal and noise: Revisiting the mitochondrial phylogeny and phylogeography of the Eurasian blindsnake species complex (Typhlopidae, Squamata). *Zoologica Scripta* 46: 665–674. <https://doi.org/10.1111/zsc.12243>
- Kotsakiozi P, Jablonski D, Ilgaz Ç, Kumlutaş Y, Avcı A, Meiri S, Itescu Y, Kukushkin O, Gvoždík V, Scillitani G, Roussos SA, Jandzik D, Kasapidis P, Lymberakis P, Poulakakis N (2018) Multilocus phylogeny and coalescent species delimitation in Kotschy's gecko,

- Mediodactylus kotschyi*: Hidden diversity and cryptic species. *Molecular Phylogenetics and Evolution* 125: 177–187. <https://doi.org/10.1016/j.ympev.2018.03.022>
- Kreiner G (2007) *Schlangen Europas*. Edition Chimaira, Frankfurt am Main, 317 pp.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35: 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Куранова ВН, Симонов ЕП, Ярцев ВВ, Шамгунова РР, Стариков ВП [Kuranova VN, Simonov EP, Yartsev VV, Shamgunova RR, Starikov VP] (2010) Разнообразие, распространение и природоохранный статус пресмыкающихся Западной Сибири. In: Дуйсебаева ТН, Зима ЮА, Чирикова МА [Duysebaeva TN, Zima YuA, Chirikova MA] (Eds) Герпетологические исследования в Казахстане и сопредельных странах. Казахстанская ассоциация сохранения биоразнообразия, Союз охраны птиц Казахстана, Алматы, 118–149.
- Kyriazi P, Kornilios P, Nagy ZT, Poulakakis N, Kumlutaş Y, Ilgaz Ç, Avcı A, Göçmen B, Lymberakis P (2013) Comparative phylogeography reveals distinct colonization patterns of Cretan snakes. *Journal of Biogeography* 40: 1143–1155. <https://doi.org/10.1111/jbi.12057>
- Litvinchuk SN, Kuranova VN, Kazakov VI, Schepina NA (2013) A northernmost record of the grass snake (*Natrix natrix*) in the Baikal Lake Region, Siberia. *Russian Journal of Herpetology* 20: 43–50.
- Madsen T (1984) Movements, home range size and habitat use of radio-tracked grass snakes (*Natrix natrix*) in southern Sweden. *Copeia* 1984: 707–713. <https://doi.org/10.2307/1445153>
- Mebert K, Trapp B, Dall'Asta A, Velenský P, Böhme W (2011) Hybrids between *Natrix tessellata* and *N. natrix/maura*. *Mertensiella* 18: 154–156.
- Meister B, Ursenbacher S, Baur B (2012) Grass snake population differentiation over different geographic scales. *Herpetologica* 68: 134–145. <https://doi.org/10.1655/herpetologica-D-11-00036.1>
- Mertens R, Wermuth H (1960) Die Amphibien und Reptilien Europas (Dritte Liste, nach dem Stand vom 1. Januar 1960). Waldemar Kramer, Frankfurt am Main, XI + 264 pp.
- Мильто КД [Milito KD] (2003) О распространении обыкновенного ужа (*Natrix natrix*) на севере европейской части России. *Современная Герпетология* 2: 100–123.
- Nielsen EE, Bach LA, Kotlicki P (2006) HYBRIDLAB (version 1.0): A program for generating simulated hybrids from population samples. *Molecular Ecology Notes* 6: 971–973. <https://doi.org/10.1111/j.1471-8286.2006.01433.x>
- Орлов НЛ, Туниев БС [Orlov NL, Tuniyev BS] (1987) Новый вид ужа *Natrix megalcephala* sp. nov. с Кавказа (Ophidia: Colubridae). Труды Зоологического Института АН СССР 158: 116–130, pl. VII.
- Orlov NL, Tuniyev BS (1992) A new species of grass snake, *Natrix megalcephala*, from the Caucasus (Ophidia: Colubridae). *Asiatic Herpetological Research* 4: 42–54.
- Pfennig KS (2021) Biased hybridization and its impact on adaptive introgression. *Trends in Ecology & Evolution* 36: 488–497. <https://doi.org/10.1016/j.tree.2021.02.010>
- Plötner J, Baier F, Akin C, Mazepa G, Schreiber R, Beerli P, Litvinchuk SN, Bilgin CC, Borkin L, Uzzell T (2012) Genetic data reveal that water frogs of Cyprus (genus *Pelophylax*) are an endemic species of Messinian origin. *Zoosystematics and Evolution* 88: 261–283. <https://doi.org/10.1002/zoos.201200021>
- Pokrant F, Kindler C, Ivanov M, Cheylan M, Geniez P, Böhme W, Fritz U (2016) Integrative taxonomy provides evidence for the species status of the Ibero-Maghrebian grass snake *Natrix astreptophora*. *Biological Journal of the Linnean Society* 118: 873–888. <https://doi.org/10.1111/bij.12782>
- Pokrant F, Kindler C, Vamberger M, Smith KT, Fritz U (2017) Grass snakes (*Natrix natrix*, *N. astreptophora*) mimicking cobras display a ‘fossil behavior.’ *Vertebrate Zoology* 67: 261–269.
- Poulakakis N, Kapli P, Karamaki A, Skortanioti E, Göçmen B, Ilgaz Ç, Kumlutaş Y, Avcı A, Lymberakis P (2013) Comparative phylogeography of six herpetofauna species in Cyprus: Late Miocene to Pleistocene colonization routes. *Biological Journal of the Linnean Society* 108: 619–635. <https://doi.org/10.1111/j.1095-8312.2012.02039.x>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Puechmaile SJ (2016) The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources* 16: 608–627. <https://doi.org/10.1111/1755-0998.12512>
- Pujolar JM, Jacobsen MW, Als TD, Frydenberg J, Magnussen E, Jónsson B, Jiang X, Cheng L, Bekkevold D, Maes GE, Bernatchez L, Hansen MM (2014) Assessing patterns of hybridization between North Atlantic eels using diagnostic single-nucleotide polymorphisms. *Heredity* 112: 627–637. <https://doi.org/10.1038/hdy.2013.145>
- Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology* 17: 285–293. <https://doi.org/10.1111/j.1365-294X.2007.03417.x>
- Rato C, Carranza S, Harris DJ (2011) When selection deceives phylogeographic interpretation: The case of the Mediterranean house gecko, *Hemidactylus turcicus* (Linnaeus, 1758). *Molecular Phylogenetics and Evolution* 58: 365–373. <https://doi.org/10.1016/j.ympev.2010.12.004>
- Reading CJ, Jofré GM (2009) Habitat selection and range size of grass snakes *Natrix natrix* in an agricultural landscape in southern England. *Amphibia-Reptilia* 30: 379–388. <https://doi.org/10.1163/156853809788795164>
- Sanz N, Araguas RM, Fernández R, Vera M, García-Marín J-L (2009) Efficiency of markers and methods for detecting hybrids and introgression in stocked populations. *Conservation Genetics* 10: 225–236. <https://doi.org/10.1007/s10592-008-9550-0>
- Schätti B (1999) *Natrix maura* (Linnaeus, 1758) – Vipernatter. In: Böhme W (Ed.) *Handbuch der Reptilien und Amphibien Europas*. Band 3/IIA: Schlangen II. Aula-Verlag, Wiebelsheim, 483–503.
- Schmitt T (2020) *Molekulare Biogeographie*. UTB, Haupt Verlag, Bern, 504 pp.
- Schultze N, Laufer H, Kindler C, Fritz U (2019) Distribution and hybridisation of barred and common grass snakes (*Natrix helvetica*, *N. natrix*) in Baden-Württemberg, South-western Germany. *Herpetozoa* 32: 229–236. <https://doi.org/10.3897/herpetozoa.32.e38897>
- Schultze N, Spitzweg C, Corti C, Delaugerre M, Di Nicola MR, Geniez P, Lapini L, Liuzzi C, Lunghi E, Novarini N, Picariello O, Razzetti E, Sperone E, Stellati L, Vignoli L, Aszталos M, Kindler C, Fritz U (2020) Mitochondrial ghost lineages blur phylogeography and taxonomy of *Natrix helvetica* and *N. natrix* in Italy and Corsica. *Zoologica Scripta* 49: 395–411. <https://doi.org/10.1111/zsc.12417>
- Sindaco R, Venchi A, Grieco C (2013) The Reptiles of the Western Palearctic, Volume 2: Annotated Checklist and Distributional Atlas of the Snakes of Europe, North Africa, Middle East and Central Asia, with an Update to Volume 1. Edizioni Belvedere, Latina, 543 pp.
- Šmíd J, Göçmen B, Crochet P-A, Trape J-F, Mazuch T, Uvizl M, Nagy ZT (2019) Ancient diversification, biogeography, and the role of

- climatic niche evolution in the Old World cat snakes (Colubridae, *Telescopus*). *Molecular Phylogenetics and Evolution* 134: 35–49. <https://doi.org/10.1016/j.ympev.2019.01.015>
- Speybroeck J, Beukema W, Bok B, van der Voort J (2016) *Field Guide to the Amphibians and Reptiles of Britain and Europe*. Bloomsbury, London, 432 pp.
- Speybroeck J, Beukema W, Dufresnes C, Fritz U, Jablonski D, Lymberakis P, Martínez-Solano I, Razzetti E, Vamberger M, Vences M, Vörös J, Crochet P-A (2020) Species list of the European herpetofauna – 2020 update by the Taxonomic Committee of the Societas Europaea Herpetologica. *Amphibia-Reptilia* 41: 139–189. <https://doi.org/10.1163/15685381-bja10010>
- Tamar K, Carranza S, in den Bosch H, Sindaco R, Moravec J, Meiri S (2015) Hidden relationships and genetic diversity: Molecular phylogeny and phylogeography of the Levantine lizards of the genus *Phoenicolacerta* (Squamata: Lacertidae). *Molecular Phylogenetics and Evolution* 91: 86–97. <http://dx.doi.org/10.1016/j.ympev.2015.05.002>
- Taylor SA, Larson EK (2019) Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology Evolution* 3: 170–177. <https://doi.org/10.1038/s41559-018-0777-y>
- Utiger U, Schätti B (2004) Morphology and phylogenetic relationships of the Cyprus racer, *Hierophis cypriensis*, and the systematic status of *Coluber gemonensis gyarosensis* Mertens (Reptilia: Squamata: Colubrinae). *Revue suisse de Zoologie* 111: 225–238. <https://doi.org/10.5962/bhl.part.80236>
- Vamberger M, Stuckas H, Vargas-Ramírez M, Kehlmaier C, Ayaz D, Aloufi AA, Lymberakis P, Široký P, Fritz U (2017) Unexpected hybridization patterns in Near Eastern terrapins (*Mauremys caspica*, *M. rivulata*) indicate ancient gene flow across the Fertile Crescent. *Zoologica Scripta* 46: 401–413. <https://doi.org/10.1111/zsc.12227>
- van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- Wisler C, Hofer U, Arlettaz R (2008) Snakes and monocultures: Habitat selection and movements of female grass snakes (*Natrix natrix* L.) in an agricultural landscape. *Journal of Herpetology* 42: 337–346. <https://doi.org/10.1670/07-027.1>
- Zhao E-M, Adler K (1993) *Herpetology of China*. Society for the Study of Amphibians and Reptiles, St. Louis, MO, 522 pp.
- Zotos S, Stamatiou M, Naziri A, Meletiou S, Demosthenous S, Perikleous K, Erotokritou E, Xenophontos M, Zavrou D, Michael K, Sergides L (2021) New evidence on the distribution of the highly endangered *Natrix natrix cypriaca* and implications for its conservation. *Animals* 11: 1077. <https://doi.org/10.3390/ani11041077>

Supplementary material 1

Table S1

Authors: Asztalos M, Ayaz D, Bayrakçı Y, Afsar M, Tok CV, Kindler C, Jablonski D, Fritz U (2021)

Data type: .xlsx

Explanation note: Table S1.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/vz.71.e76453.suppl1>

Supplementary material 2

Supplementary Figures S1–S3

Authors: Asztalos M, Ayaz D, Bayrakçı Y, Afsar M, Tok CV, Kindler C, Jablonski D, Fritz U (2021)

Data type: .pdf

Explanation note: Supplementary Figures S1–S3, Supplementary Tables S2–S11.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/vz.71.e76453.suppl2>