VERTEBRATE ZOOLOGY

68 (3): 269–281
20.11.2018

SENCKENBERG

Phylogeography and taxonomy of the barred grass snake (*Natrix helvetica*), with a discussion of the subspecies category in zoology

Carolin Kindler & Uwe Fritz

Museum of Zoology, Senckenberg Dresden, A. B. Meyer Building, 01109 Dresden, Germany; caro.kindler@freenet.de; uwe.fritz@senckenberg.de

Accepted October 15, 2018.
Published in print and online at www.senckenberg.de/vertebrate-zoology on November 20, 2018.

Editor in charge: Martin Päckert

Abstract

Using 12 microsatellite loci, mitochondrial DNA sequences and previously published morphological evidence, we examined the phylogeographic and taxonomic structure of *Natrix helvetica*. Our results support tentatively the recognition of five subspecies: (1) *N. h. helvetica* (Lacepède, 1789) from Western Europe, (2) *N. h. cetti* Gené, 1839 from Sardinia, (3) *N. h. corsa* (Hecht, 1930) from Corsica, (4) *N. h. lanzai* Kramer, 1970 from the Po drainage and peninsular Italy, except for parts of Calabria and perhaps Apulia, and (5) *N. h. sicula* (Cuvier, 1829) from Sicily and parts of Calabria. The status of the subspecies from Corsica and Sardinia warrants further research. Grass snakes from the two islands are genetically deeply divergent from their continental and Sicilian conspecifics. However, the mitochondrial haplotypes of *N. h. cetti* and *N. h. corsa* are not reciprocally monophyletic, and our analyses of microsatellite data did not unambiguously support their distinctness. *Natrix helvetica lanzai* comprises three mitochondrial lineages and could represent more than one taxon. Grass snakes of the northern mitochondrial lineage of *N. h. lanzai* (mainly from the Po drainage system) were previously assigned to *N. h. helvetica*. We propose that the recognition of subspecies within *N. helvetica* reflects its genetic differentiation best. Our subspecies concept resembles the definition of Evolutionarily Significant Units (ESUs) in that subspecies should be ideally confirmed by two independent genetic lines of evidence (mtDNA, nuclear genomic markers). In contrast to species, subspecies are still fully capable of extensive gene flow and may disappear in the evolutionary process as a consequence of secondary contact. Thus, subspecies represent an early stage of (incomplete) speciation. Using the subspecies category facilitates communication within and beyond science (legislation, conservation), a clear advantage compared to the cumbersome concept of ESUs.

Key words

Biogeography, Corsica, Evolutionarily Significant Units, Italy, Sardinia, speciation, species, Western Europe.

Introduction

True grass snakes are widely distributed and abundant in the Western Palearctic. All true grass snakes were traditionally placed into the polytypic species *Natrix natrix* (Linnaeus, 1758) (Mertens & Wermuth, 1960; Kabisch, 1999), even though some authors suggested the existence of two further species based on morphological differences, *Natrix cetti* Gené, 1839 and *Natrix megalocephala* Orlov & Tuniyev, 1987 (Orlov & Tuniyev, 1987; Vanni & Cimmaruta, 2011). However, these two putative species were not confirmed by molecular markers and their validity was rejected (Fritz *et al*., 2012; Kindler *et al*., 2013), so that the genus *Natrix* comprised until recently only three species, the grass snake *N. natrix*, the viperine snake *N. maura* (Linnaeus, 1758) and the dice snake *N. tessellata* (Laurenti, 1768). Grass snakes were thought to inhabit the largest distribution range of all, from Northwestern Africa and the Iberian Peninsula through most of Europe eastwards to Central Asia. Based on morphology, up to 14 subspecies of *N. natrix* were distinguished (Kabisch, 1999). Several studies using mitochondrial DNA (mtDNA) sequences and microsatellite loci (Guicking *et al*., 2006, 2008; Fritz *et al*., 2012; Kindler *et al*., 2013, 2014, 2017, 2018a, b; Pokrant *et al*., 2016) elucidated

Table 1. Currently recognized subspecies of *Natrix helvetica*, their distribution ranges and mitochondrial lineages (combined from Kabisch, 1999; Kindler *et al.*, 2013, 2017). In Calabria is a contact zone of *N. h. lanzai* and *N. h. sicula* (Kindler *et al.*, 2013).

Subspecies	Distribution	mtDNA lineage (Kindler *et al*., 2013)
N. h. helvetica (Lacepède, 1789)	Rhine region westwards to the Pyrenees, Britain, Po Plain	C, E
N. h. cetti Gené, 1839	Sardinia	B
N. h. corsa (Hecht, 1930)	Corsica	B
N. h. lanzai Kramer, 1970	Apennine Peninsula south of Po Plain, except Calabria	D, F
N. h. sicula (Cuvier, 1829)	Sicily, Calabria	A

the genetic structure of grass snakes, resulting in the recent recognition of three distinct species (Pokrant *et al*., 2016; Kindler *et al*., 2017), none of which corresponds to those two species proposed on morphological grounds. The three currently accepted true grass snake species are, from west to east, the Ibero-Maghrebian red-eyed grass snake *N. astreptophora* (Seoane, 1884); the barred grass snake *N. helvetica* (Lacepède, 1789), distributed from the Pyrenees eastwards to the Rhine region, including Britain, the Apennine Peninsula, Sicily, Corsica, and Sardinia; and the common or eastern grass snake *N. natrix*, occurring from the Rhine region eastwards to Central Asia.

The barred grass snake comprises several morphologically distinctive forms that have been traditionally recognized as distinct subspecies (Mertens & Wermuth, 1960; Kabisch, 1999; Table 1). The widest distribution has the nominotypical subspecies *N. h. helvetica*, which occurs in Britain and from the Pyrenees to the Rhine region; all other subspecies are more localized. Among the other subspecies are the geographically isolated and morphologically distinctive grass snakes from Corsica and Sardinia, which once have been thought to be morphologically so divergent that they deserve recognition as two 'incipient species' in the rank of subspecies of *N. natrix* (Thorpe, 1979) or as one distinct species with two subspecies (*N. cetti cetti* and *N. c. corsa*; Vanni & Cimmaruta, 2011). While the contact zone of *N. helvetica* and *N. natrix* has been studied in detail using mtDNA and 13 microsatellite loci (Kindler *et al*., 2017), genetic variation within *N. helvetica* has been examined until now only using mtDNA. Based on this marker, *N. helvetica* is comprised of six lineages which are successive sister taxa in phylogenetic analyses. These lineages do not completely match with subspecies (Table 1). Unexpectedly, grass snakes from Corsica and Sardinia were neither the most divergent lineage nor reciprocally monophyletic. Instead, the mitochondrial lineage of Sicilian and Calabrian grass snakes, corresponding to the subspecies *N. h. sicula*, was sister to a clade containing the other five lineages, including that from Corsica and Sardinia (Kindler *et al*., 2013). According to a fossil-calibrated molecular clock (Kindler *et al*., 2018b), mitochondrial divergence within *N. helvetica* commenced approximately 6.8 million years ago (mya), with the lineage of *N. h. sicula* being the oldest. The Corso-Sardinian lineage came out as sister to the three lineages from the Apennine Peninsula and the Padan Plain plus the widely distributed lineage E that corresponds to *N. h. helvetica* (Kindler *et al*., 2013). The Corso-Sardinian lineage was estimated to have diverged approximately 3.9 mya, whereas the remaining lineages were dated to 2.75 to 0.3 mya, with the split between lineage E of *N. h. helvetica* and lineage F of *N. h. lanzai* being the youngest (Kindler *et al*., 2018b).

The wide distribution of lineage E is understood as the result of a rapid Holocene range expansion from a glacial refuge in Southern France, in accordance with weak differentiation of mitochondrial haplotypes (Kindler *et al*., 2018a). However, mtDNA is exclusively inherited maternally. Therefore, in bisexual species, nuclear genomic differentiation may differ significantly. To examine whether this could be also the case in barred grass snakes, we used range-wide sampling (Fig. 1) and nuclear microsatellite loci, combined with evidence from mtDNA. Our objective was to address the following questions: (1) Do the mitochondrial lineages of *N. helvetica* match or do they conflict with nuclear genomic clusters, and (2) is the widely distributed nominotypical subspecies differentiated in local population clusters that are not mirrored by mtDNA? If the mitochondrial lineages match with nuclear genomic clusters, this would support taxonomic distinctness of the concerned populations. In addition, a finer population structuring of *N. h. helvetica* would have implications for conservation and could necessitate the recognition of distinct Management Units in the sense of Moritz (1994). Finally, we used our results for a critical discussion of the subspecies category in zoology.

Materials and Methods

The same genetic markers were employed as in our previous studies on grass snakes (Kindler *et al*., 2013, 2014, 2017, 2018a, b; Pokrant *et al*., 2016), i.e. 13 unlinked polymorphic microsatellite loci (Pokrant *et al*., 2016; Kindler *et al*., 2017, 2018b) and two mtDNA fragments, the partial ND4 gene plus adjacent DNA coding for tRNAs (867 bp) and the cyt *b* gene (1,117 bp). The majority of data was already used for our previous publications (Kindler *et al*., 2013, 2014, 2017, 2018a, b; Pokrant *et al*., 2016) but for other purposes. Data for seven samples from Italy were new (Table S1) and produced accord-

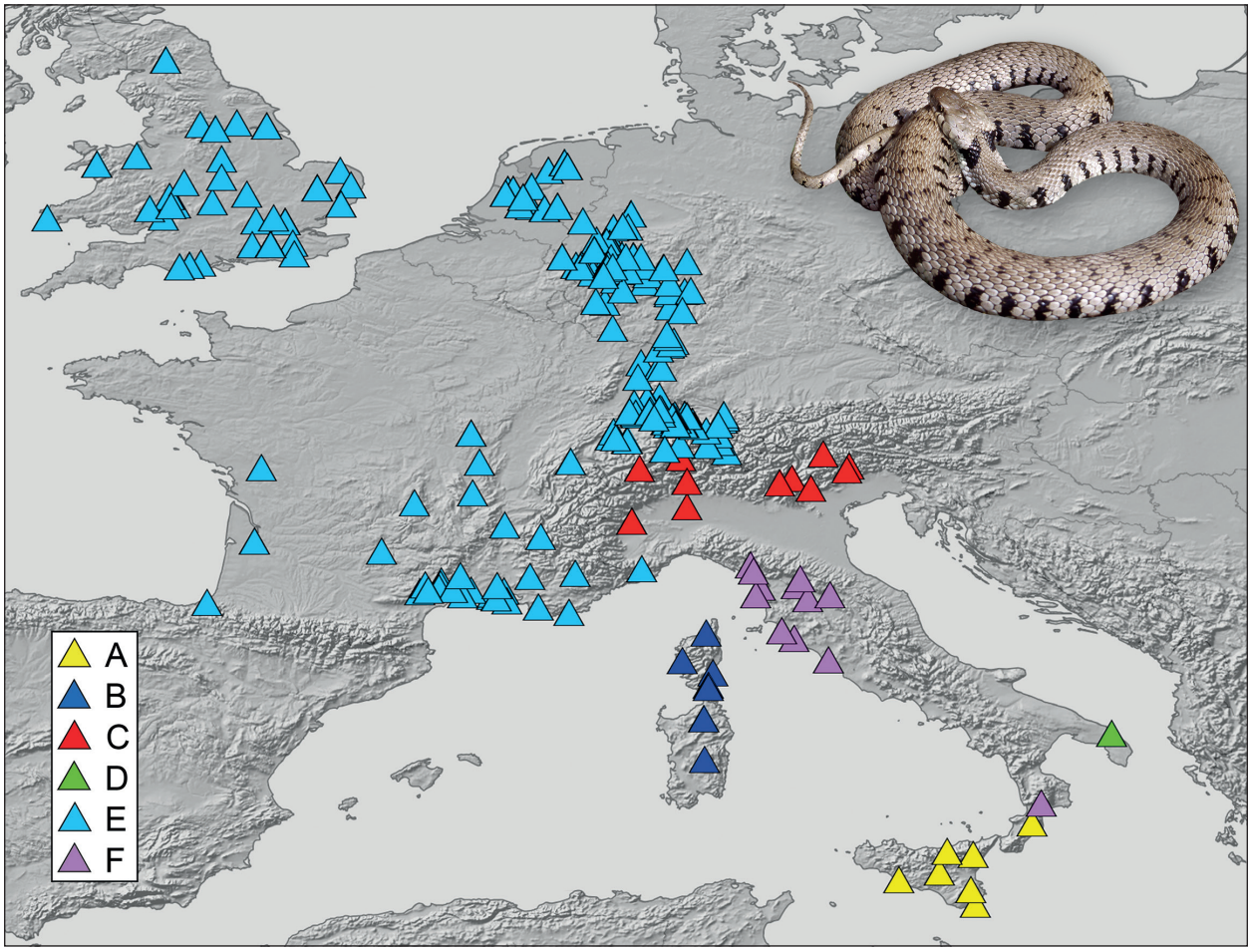

Fig. 1. Sampling sites and mitochondrial lineages of *Natrix helvetica* (*n*=450 individuals). Inset: *N. h. helvetica* (Linz am Rhein, Germany); photo: Wolfgang Böhme.

ing to laboratory procedures described in Kindler *et al*. (2013) and Pokrant *et al*. (2016). In total, data from 450 grass snakes were analyzed (Fig. 1; Table S1); 408 samples represented mitochondrial lineage E; seven samples corresponded to lineage A; nine samples to lineage B, 12 samples to lineage C, and 13 samples to lineage F. Lineage D was represented only by GenBank data for the mtDNA of one snake (Guicking *et al*., 2006). For analyses of microsatellites, 74 samples of lineage E of *Natrix helvetica helvetica* that showed evidence for hybridization with *N. natrix* in Kindler *et al*. (2017) were excluded. Names of mitochondrial lineages follow throughout the paper Kindler *et al*. (2013). Accession numbers for haplotypes are given in Table S2.

Mitochondrial DNA sequences were aligned using bioedit 7.0.9.0 (Hall, 1999), resulting in an 867-bp-long alignment of 442 sequences corresponding to the ND4 gene and adjacent DNA and an 1,117-bp-long alignment of 351 cyt *b* sequences (four cyt *b* sequences from Kindler *et al*., 2013 were excluded because they were too short for network calculations). For some samples only one mtDNA fragment was available (Table S1). The mitochondrial lineage of each new sample was identified by running exploratory Maximum Likelihood (ML) analyses using RAxML 7.2.8 (Stamatakis, 2006) including previously published data (Kindler *et al*., 2013, 2017), the GTR + G model and a fast ML search with 100 bootstrap values. Then, a parsimony network was drawn for each alignment using TCS (Clement *et al*., 2000), with gaps coded as fifth character state. Each fragment was analyzed separately because TCS software cannot cope with missing data. Using the default 95% connection limit, unconnected haplotype clusters were obtained for the different genetic lineages, which is why the connection limit was arbitrarily set to 50 steps. Based on haplotypes, uncorrected *p* distances (means) were calculated using MEGA 10.0.1 (Kumar *et al*., 2018) and the pairwise deletion option.

For examining the genotypic structure of *Natrix helvetica*, the same 13 microsatellite loci as in Kindler *et al*. (2017) were used. However, one locus (Nsμ3) was excluded from calculations due to missing data. The other 12 loci were examined for Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium using arlequin 3.5.1.3 (Excoffier & Lischer, 2010). The presence of null alleles was tested in micro-checker 2.2.3 (van Oosterhout *et al*., 2004). No evidence was found for null alleles, linkage disequilibrium or deviation from HWE. Microsatellite data were then subjected to unsupervised cluster analyses using structure 2.3.4 (Pritchard *et al*.,

2000; FALUSH *et al*., 2003). STRUCTURE searches in the data set for partitions that are, as far as possible, in HWE and, within each partition, in linkage equilibrium between loci (PRITCHARD *et al*., 2000). Because STRUCTURE is known to be sensitive against uneven sample sizes (PUECHMAILLE, 2016), data for the nominotypical subspecies were rarefied by randomly drawing a subsample of 20 individuals (10 from Southern France and adjacent Italy plus 10 from the remaining distribution range). The resulting data set ($n=49$) consisted of all available data for southern barred grass snakes plus the data for 20 *N. h. helvetica* and was analyzed with STRUCTURE using correlated allele frequencies and the admixture model. The Monte Carlo Markov chains ran for 1 million generations, and the burn-in was set to 250,000 generations. For *K*s ranging from 1 to 10, calculations were repeated ten times each. The best *K* value was determined using STRUCTURE HARVESTER (EARL & VONHOLDT, 2012). In addition, posterior probabilities for *K* values were inspected. STRUCTURE results were finally visualized using DISTRUCT 1.1 (ROSENBERG, 2004). For the microsatellite data of each STRUCTURE cluster and the corresponding mtDNA data, pairwise F_{ST} values and Analyses of Molecular Variance (AMOVAs) were obtained using ARLEQUIN. To examine the genetic distinctiveness of the clusters without underlying population genetic presumptions, Principal Component Analyses (PCAs) were calculated using the R package ADEGENET (JOMBART, 2008). Furthermore, to examine the geographic structuring of *N. h. helvetica*, microsatellite data of all 334 samples of this subspecies were subjected to STRUCTURE analyses and PCAs, employing for STRUCTURE the same settings as described above.

Results

Parsimony networks and uncorrected p distances of mtDNA, geographic distribution of mtDNA lineages. Each previously identified mtDNA lineage (KINDLER *et al*., 2013) corresponded to a distinct haplotype cluster in parsimony network analysis, both for the mtDNA fragment consisting of the partial ND4 gene plus adjacent DNA (ND4 + tRNAs) and for the cyt *b* gene.

Records of lineage A matched with all samples from Sicily and one sample from Calabria, where lineage F was recorded in close proximity. Lineage B was represented by all samples from Corsica and Sardinia. Records for lineage C were mainly from the catchment basin of the Po River, with records from South Tyrol, Trentino, Piedmont, and Veneto in Italy, and within Switzerland, from Ticino and Valais (Wallis). A record from Lago Ritóm, Ticino, Switzerland is from 1,900 m a.s.l.. Another Swiss record was from beyond the Simplon Pass (2,005 m a.s.l.), outside the Po drainage (Niedergesteln, Valais). Lineage D was represented only by GenBank sequences from a single individual from Southern Apulia (GUICKING *et al*., 2006). Lineage E was recorded from continental France, Great Britain, the Netherlands, Western Germany, Switzerland, and a single site in Northwestern Italy, at the Piedmontese-Ligurian border close to France. Lineage F corresponded to samples from mainland Italy (Tuscany, Lazio, Calabria).

For ND4 + tRNAs (Fig. 2, top), there was a star-like cluster of 12 haplotypes corresponding to lineage E. These haplotypes differed by a maximum of two mutation steps. The unique haplotype of the Apulian lineage D, identified from one GenBank sequence only (GUICKING *et al*., 2006), was connected to cluster E by six mutation steps. The Corso-Sardinian cluster B comprised two haplotypes separated by one mutation step, and was connected to cluster E by a minimum of 20 mutation steps. Samples from Corsica and Sardinia had shared haplotypes (Table S1). Cluster A from Sicily and Calabria consisted of three haplotypes that differed by a maximum of two mutations. This cluster was connected by 29 mutation steps to the Corso-Sardinian cluster and by 45 steps to cluster E. In cluster F, from western mainland Italy, occurred two haplotypes with seven mutation steps difference that were connected by a minimum of two steps with cluster E. Lineage C from Northern Italy had only one haplotype that differed by a minimum of 18 steps each from lineages F and E.

For the cyt *b* gene (Fig. 2, bottom) a similar but more complex pattern was revealed, again with the cluster of haplotypes from Sicily (A) being the most distinct. Lineage E was represented by a cluster of 29 haplotypes, with individual haplotypes differing by a maximum of five mutation steps. The remaining clusters were connected over a large loop. This loop diverged from cluster E via three mutations that led to cluster F. Cluster F itself consisted of five haplotypes that were connected over another small loop by a maximum of five steps. Cluster F was linked over at least 33 steps with cluster B. Cluster B had five haplotypes differing by a maximum of five mutations, and no shared haplotypes were recorded for Corsica and Sardinia (Table S1). Cluster B was connected over at least 57 steps with cluster A, and over a minimum of 33 steps with cluster C. In cluster A occurred four haplotypes having a maximum of four mutations. Cluster A and cluster C were connected by a minimum of 50 steps. Cluster C had two haplotypes differing by three mutations. These two haplotypes were connected over at least 18 steps to a missing node haplotype leading over four additional steps to the unique haplotype of lineage D, and over another six steps to cluster F to close the loop.

Based on haplotypes, the six lineages differed for ND4 + tRNAs by uncorrected *p* distances ranging from 0.63% to 6.01% (within lineage divergences: 0.12–0.81%) and, for the cyt *b* gene, by 0.45% to 5.22% (within lineage divergences: 0.13–0.27%; Table S3).

Microsatellite analyses and comparisons with mtDNA differentiation. For the data set containing 20 randomly selected *Natrix helvetica helvetica* and all southern samples STRUCTURE HARVESTER suggested $K=2$ as optimal number of clusters. However, the ΔK value for $K=3$ was very similar (Fig. S1). For $K=2$, all samples

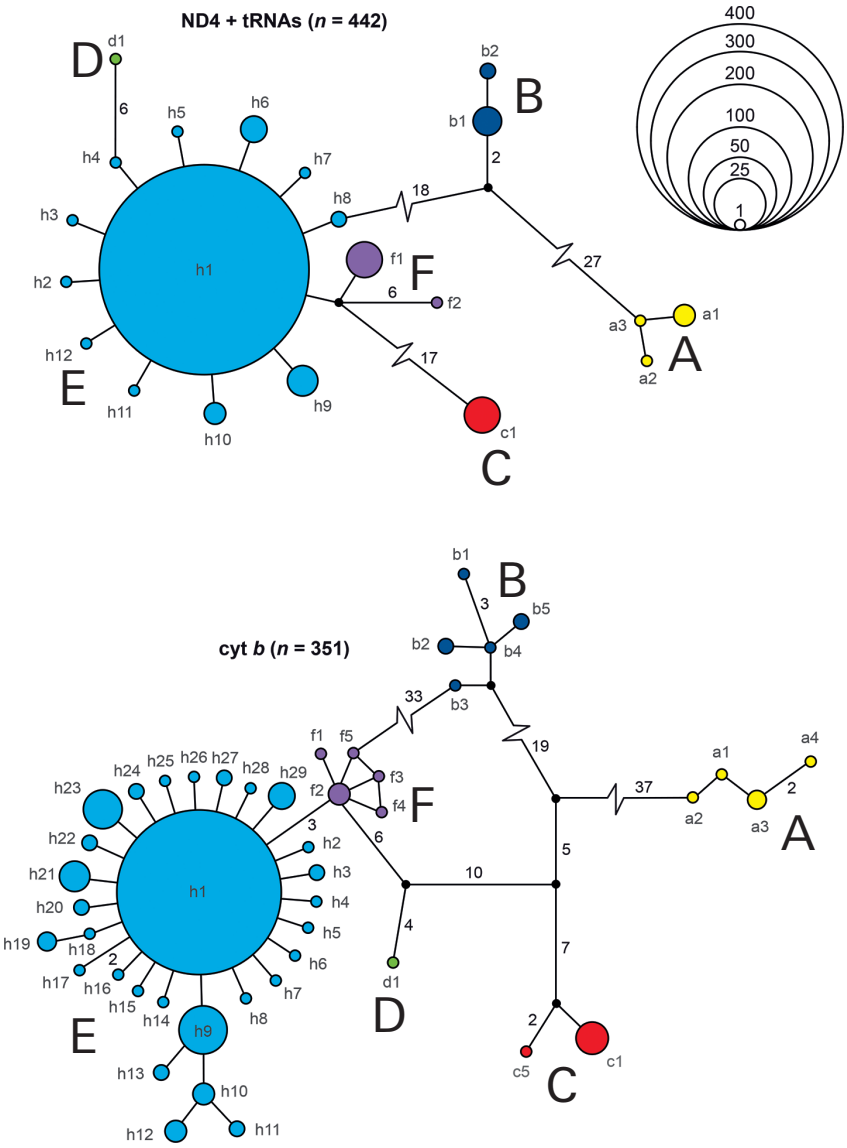

Fig. 2. Parsimony networks for mtDNA sequences of *Natrix helvetica*. Symbol size corresponds to haplotype frequency. Haplotype colors correspond to mitochondrial lineages as in Figure 1. Small black circles are missing node haplotypes. Lines connecting haplotypes represent one mutation step, if not otherwise indicated by numbers of substitutions along lines. Individual haplotype names in grey; for European Nucleotide Archive (ENA) accession numbers, see Table S2.

of *N. h. helvetica* corresponded to one cluster and all southern samples to the other (Fig. 3, top). Admixture between the two clusters occurred in Southern France and Northwestern Italy (Fig. 4, left). With respect to the mitochondrial lineages, there was a perfect agreement between lineage E of KINDLER *et al*. (2013) and the nuclear cluster of *N. h. helvetica* (Fig. 3). The second nuclear cluster included all samples from Italy and Corsica, which harbored four distinct mitochondrial lineages (A, B, C, F). According to an AMOVA using microsatellite data, 85.09% of the molecular variance occurred within and 14.91% between these two clusters, corresponding to an F_{ST} value of 0.15.

If the mtDNA sequences of the same clusters were compared in AMOVAs, for the sequences containing mtDNA coding for ND4 + tRNAs, 66.67% of the molecular variance occurred within and 33.33% between the two clusters (F_{ST} = 0.33). For the cyt *b* gene, 65.54% of the molecular variance occurred within and 34.46% between the clusters (F_{ST} = 0.35). When the AMOVAs were calculated using the five mitochondrial lineages A–F as groups, the following pattern was revealed: For ND4 + tRNAs, only 2.39% of the molecular variance occurred within and 97.61% among the groups (F_{ST} = 97.61). For the cyt *b* gene, 3.16% of the molecular variance occurred within and 96.84% among and the groups (F_{ST} = 96.84).

STRUCTURE is known to reveal only the uppermost hierarchical level of genetic differentiation (EVANNO *et al*., 2005). Therefore, the data of the southern microsatellite cluster were examined alone, exclusive of two individuals with admixture proportions beyond 20%. For this second STRUCTURE analysis, K = 2 was also revealed as the best number of clusters (Fig. S2), and the snakes from Corsica and Sardinia were assigned to another cluster than those from mainland Italy and Sicily (Fig. 3, center; Fig. 4, right). In an AMOVA, 77.93% of the molecular variance of the microsatellites occurred within, and 22.07% between the two clusters, equaling an F_{ST} value of 0.22. The calculations for the complete data set using K = 3 matched perfectly with these results in that Corso-Sardinian grass snakes constituted the third cluster, while

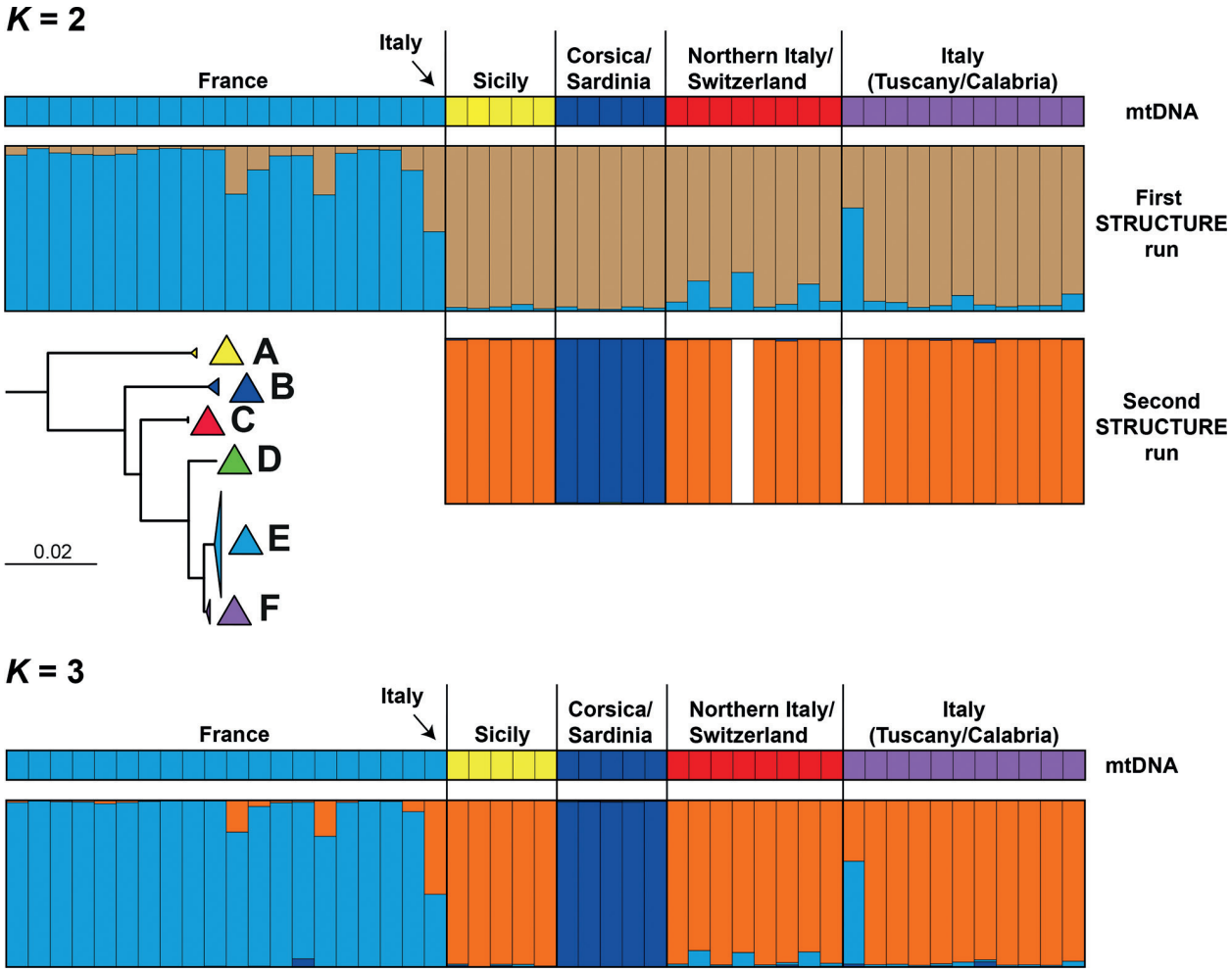

Fig. 3. Genotypic structuring of 49 *Natrix helvetica*. The mitochondrial lineage of each sample is shown above the STRUCTURE diagrams. Haplotype colors correspond to lineages: light blue (lineage E), yellow (lineage A), dark blue (lineage B), red (lineage C), green (lineage D) and lilac (lineage F). In the STRUCTURE diagrams, an individual sample is represented by a vertical bar reflecting its inferred ancestry. White bars in the second run correspond to excluded admixed snakes. Top: clustering results for $K=2$ for all samples and the southern subsample. Inset: Branching pattern of mitochondrial lineages from KINDLER *et al.* (2013). Bottom: clustering results for $K=3$ for all samples.

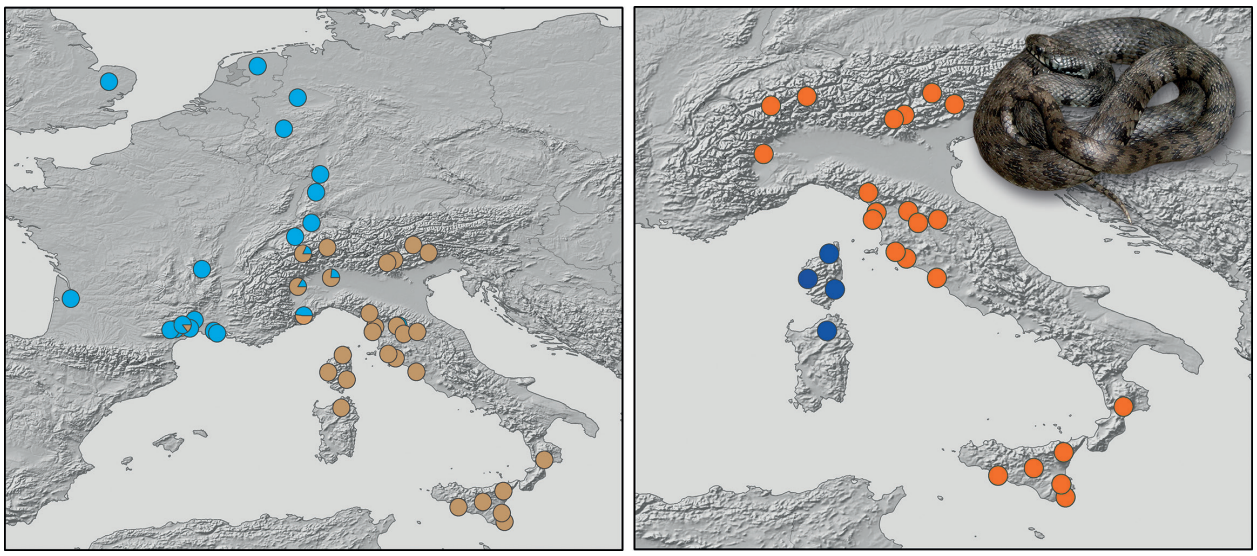

Fig. 4. Genotypic assignment of grass snakes using STRUCTURE analyses of 12 microsatellite loci. Symbol colors correspond to STRUCTURE barplots (Fig. 3). Mixed ancestries are indicated by differently colored sectors corresponding to inferred genetic percentages of the respective cluster. Left: STRUCTURE run including 20 randomly selected samples of *Natrix helvetica helvetica* and all southern samples. Right: STRUCTURE run without *N. h. helvetica*. Inset: *N. h. cetti* (Sette Fratelli, Sardinia); photo: Philippe Geniez.

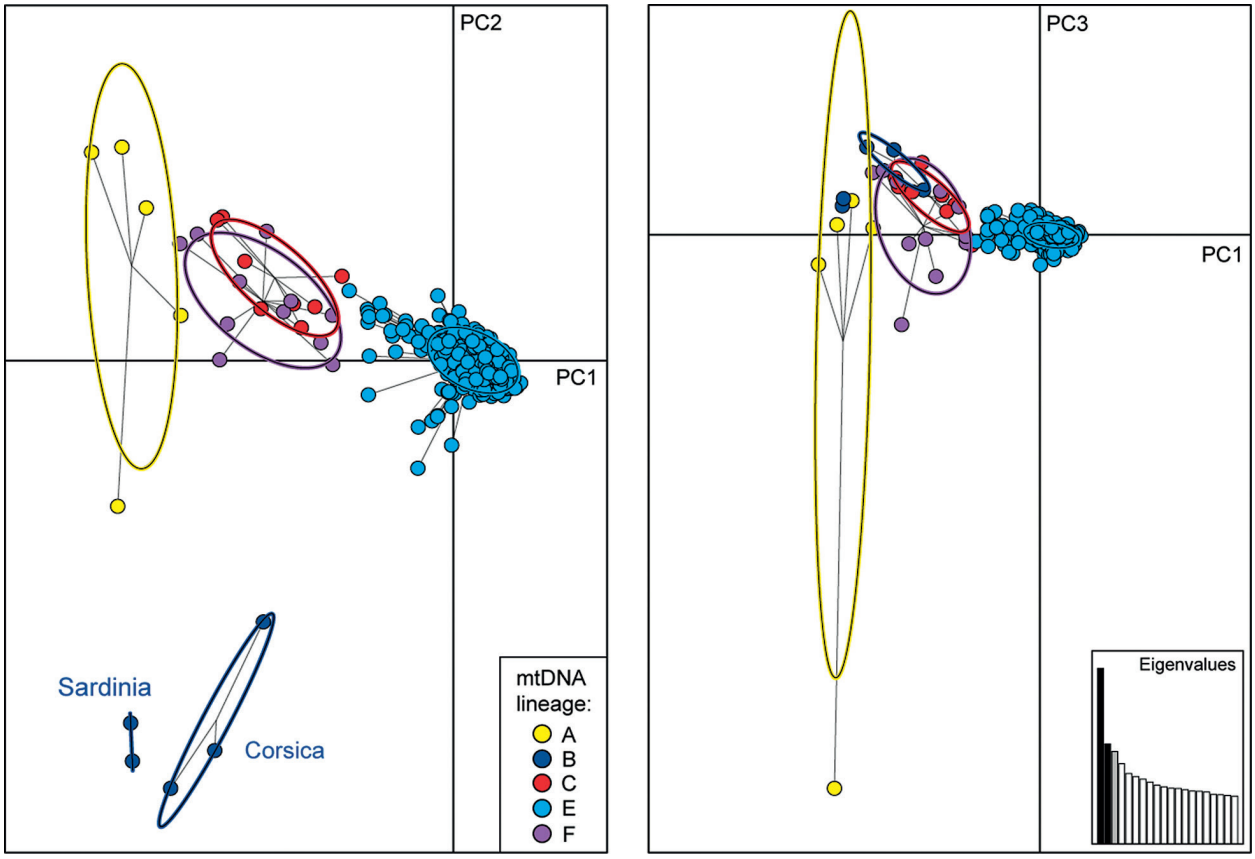

Fig. 5. PCA for microsatellite data of 363 samples of *Natrix helvetica*. Samples are colored according to mitochondrial lineages. The oval outlines represent 95% confidence intervals. The x axis (PC1) explains 7.9 % and the y axes 4.5% (PC2) and 4.2% (PC3) of the variance.

the other cluster assignments and admixture proportions remained virtually identical (Fig. 3, bottom).

AMOVAs for the mtDNA of the samples used in the second STRUCTURE analyses (Fig. 3, center) found for the sequences containing the partial ND4 gene and DNA coding for tRNAs 58.63% of the molecular variance within and 41.37% between the two clusters (F_{ST}=0.41) and for the cyt *b* gene, 52.89% within and 47.11% between the clusters (F_{ST}=0.47).

The PCAs showed much better resolution than the STRUCTURE analyses. The PCAs returned most grass snakes harboring different mitochondrial lineages as distinct clusters, and the two samples from Sardinia differed also from the three samples from Corsica (Fig. 5). Broad overlap occurred only with respect to the samples corresponding to mtDNA lineages C (Po drainage) and F (western Apennine Peninsula).

When STRUCTURE analyses were run for all 334 samples of *N. h. helvetica* alone (exclusive of all southern samples representing mtDNA lineages A–C and F), STRUCTURE HARVESTER found K=3 and K=4 as equally good solutions for the cluster numbers; K=1 could be excluded after inspection of the posterior probabilities for K (Fig. S3). However, both for K=3 and K=4 high levels of admixture were revealed, indicating weak differentiation and extensive gene flow across the range (Figs 6 and 7). This weak differentiation is in line with the results of a PCA for the 334 samples (Fig. S4).

Discussion

Using 12 microsatellite loci as nuclear genomic markers, our study found clusters resembling, but not completely matching, the differentiation pattern of the mitochondrial lineages of *Natrix helvetica* (Fig. 8). Grass snakes of the nominotypical subspecies *N. h. helvetica* match excellently with the distribution of the mitochondrial lineage E and the corresponding microsatellite cluster, except for Northern Italy (see below). The weak genetic structuring within *N. h. helvetica* (Figs 6, 7, S4) supports that its wide distribution range was rapidly colonized in the Holocene (KINDLER *et al.*, 2018a).

One of the six previously identified mitochondrial lineages of *N. helvetica* (KINDLER *et al.*, 2013) could not be studied using microsatellites (lineage D from Apulia, only known from GenBank sequences from one individual; GUICKING *et al.*, 2006). For the remaining five lineages microsatellites could be generated. Among these, lineage A from Sicily and Calabria is most divergent and sister to all remaining lineages of *N. helvetica* (Fig. 8), with an estimated mean age of 6.83 mya (KINDLER *et al.*, 2018b). This is in agreement with a general phylogeographic paradigm that Sicily harbors deeply divergent genetic lineages that are typically sister to lineages from the Apennine Peninsula and sometimes more northerly regions (as reviewed in KINDLER *et al.*, 2013).

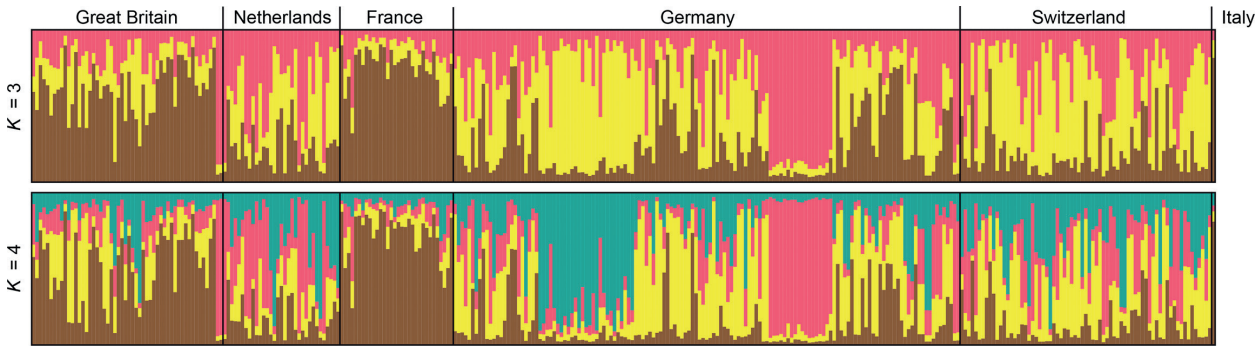

Fig. 6. Genotypic structuring of 334 *Natrix helvetica helvetica* for $K=3$ (top) and $K=4$ (bottom). The high level of genetic admixture indicates negligible differentiation and extensive gene flow. For further explanation, see Figure 3.

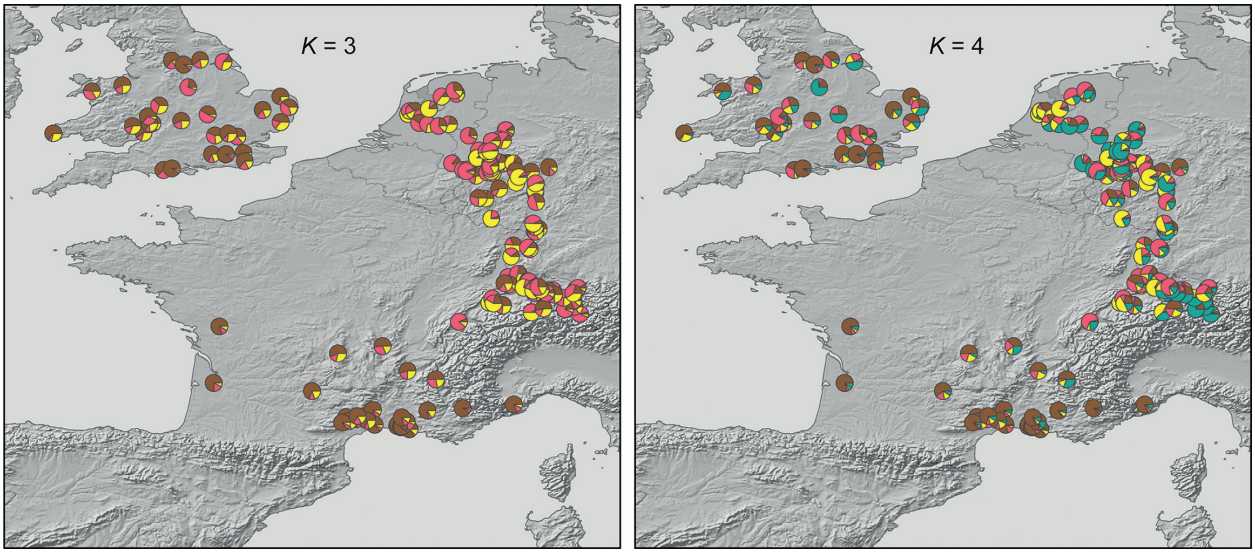

Fig. 7. Genotypic structuring of 334 *Natrix helvetica helvetica* for $K=3$ and $K=4$. Colors correspond to barplots in Figure 6.

The geographic distribution of lineage A matches the putative range of the morphologically defined subspecies *N. h. sicula* (Table 1). Our unsupervised STRUCTURE analyses did not reveal a distinct nuclear cluster for the samples representing lineage A (Fig. 3). In contrast, our PCA showed that this lineage is clearly distinct (Fig. 5). STRUCTURE is known to discover only the uppermost hierarchical level of divergence (EVANNO *et al.*, 2005). Moreover, the software is sensitive against uneven sample sizes (PUECHMAILLE, 2016) and, despite our efforts to generate similar sample sizes, we cannot exclude that the clustering results of STRUCTURE are still biased from the small and still somewhat uneven sample sizes for the individual southern lineages. In any case, STRUCTURE and PCA calculations agreed in finding lineages E and B distinct, and PCA discriminated within lineage B even samples from Corsica and Sardinia. In contrast, lineages C and F were not distinct in STRUCTURE analyses and PCA (Figs 3, 5, 8). Lineage E corresponds to the samples of *N. h. helvetica* and lineage B to the Sardinian subspecies *N. h. cetti* plus the Corsican subspecies *N. h. corsa*. Lineages C and F are currently identified with the subspecies *N. h. helvetica* and *N. h. lanzai*, respectively (Table 1).

While the morphological distinctness of *N. h. lanzai* and *N. h. sicula* has been doubted (THORPE, 1979, 1980, 1984a, b), the validity of the Corsican and Sardinian taxa has been generally accepted based on their peculiar morphology (MERTENS, 1947, 1957; MERTENS & WERMUTH, 1960; THORPE, 1979, 1980, 1984b; KABISCH, 1999; VANNI & CIMMARUTA, 2011). Using multivariate analyses of many morphological characters, THORPE (1979) regarded each of the two Corso-Sardinian taxa as an 'incipient species' in the rank of a subspecies. In addition to the Corsican and Sardinian subspecies, THORPE (1979) recognized only two additional continental grass snake subspecies in which he lumped together all other subspecies. This reduction of the number of continental subspecies to two was not unambiguously appreciated (KRAMER *et al.*, 1982; ENGELMANN *et al.*, 1986; GRUBER, 1989; BRAÑA, 1998; KABISCH, 1999; KREINER, 2007; GENIEZ, 2015), and many authors continued to distinguish additional subspecies. VANNI & CIMMARUTA (2011) placed grass snakes from Corsica and Sardinia in the distinct species *Natrix cetti*, with a Corsican (*N. c. corsa*) and a Sardinian subspecies (*N. c. cetti*). However, both the views of THORPE (1979) and VANNI & CIMMARUTA (2011) conflicted with the mitochondrial differentiation of grass snakes, and

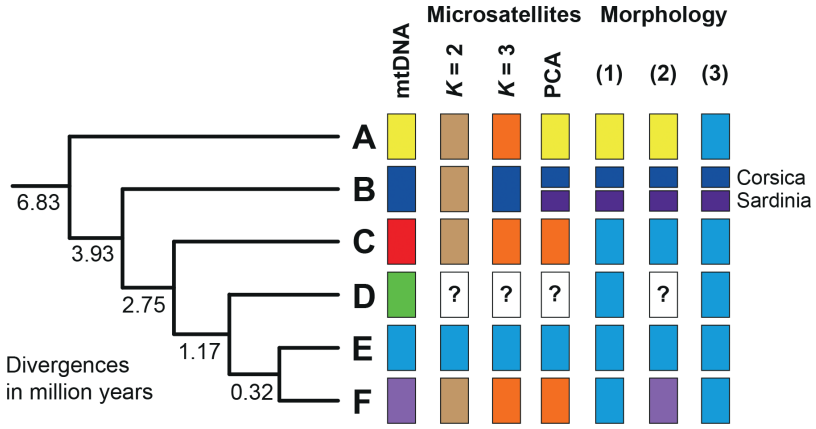

Fig. 8. Comparison of different genetic markers and morphology for *Natrix helvetica*. Phylogenetic branching pattern of mtDNA lineages according to Kindler *et al.* (2013); microsatellite analyses according to the present investigation; morphological differences according to (1) Mertens & Wermuth (1960), (2) Kramer (1970), Kramer *et al.* (1982), and (3) Thorpe (1979). The latter author included in his concept of the subspecies *helvetica* also what is now *N. astreptophora*. Numbers at nodes are estimates for mean divergence times of mitochondrial lineages from Kindler *et al.* (2018b).

the elevation of Corso-Sardinian grass snakes to species level was rejected (Fritz *et al*., 2012). Subsequent studies using mtDNA and microsatellite loci to examine the phylogeographic structuring of grass snakes (Kindler *et al*., 2013, 2014, 2017, 2018a, b; Pokrant *et al*., 2016) revealed the existence of three distinct species, each with multiple genetic lineages. The red-eyed grass snake *N. astreptophora* (Seoane, 1884) is distributed over North Africa and the Iberian Peninsula (Pokrant *et al*., 2016; Kindler *et al*., 2018b). Its range abuts in Southwestern France with that of the barred grass snake, *N. helvetica* (Lacepède, 1789). Hybridization between the two species is rare (Pokrant *et al*., 2016). *Natrix helvetica* is distributed from the Pyrenees across mainland France to the Rhine region and also occurs in Britain, Italy (including Sicily and Sardinia) and on Corsica. In a narrow strip in the Rhine region it hybridizes with the common grass snake *Natrix natrix* (Linnaeus, 1758), with largely unidirectional gene flow from *N. helvetica* into *N. natrix* (Kindler *et al*., 2017). The range of the latter species includes Fennoscandia and extends from Central Europe across Eastern and Southeastern Europe to the Middle East and Central Asia (Bannikov *et al*., 1977; Kabisch, 1999; Sindaco *et al*., 2013; Kindler *et al*., 2014).

Thorpe (1980, 1984b) regarded the morphological variation within what is now *N. astreptophora* and *N. helvetica* as clinal and suggested that two ancestral populations extended their ranges northwards, one from Southern Italy and Sicily and the other from the Iberian Peninsula and North Africa, with the two range expansions meeting in Western and Central Europe. This hypothesis is untenable in the light of the genetic differentiation of the involved grass snake populations representing two distinct biological species, *N. astreptophora* and *N. helvetica* (Pokrant *et al*., 2016). In addition, mtDNA provided unambiguous evidence for several glacial refugia of *N. helvetica* in Southern France, Corso-Sardinia, the Apennine Peninsula and Sicily (Kindler *et al*., 2013, 2018a). In the following, we will discuss whether it is justified to recognize these lineages as distinct taxa and if so, which taxonomic rank should be applied – species or subspecies.

What is a subspecies? While there is an ever-growing body of literature about different species concepts and their pros and cons (see for instance the recent review in Zachos, 2016), distinctly less attention has been paid to the subspecies category. Within the International Code of Zoological Nomenclature (ICZN, 1999), and thus in the Linnean system, the rank of subspecies is recognized in addition to the rank of species. These two ranks together constitute the so-called species group (ICZN, 1999: Article 45). For some vertebrate groups, in particular mammals (Wilson & Reeder, 2005), birds (Clements *et al*., 2018) and reptiles (Uetz *et al*., 2018), subspecies have traditionally been distinguished and are still widely used, also in legislation and conservation. A conceptual weakness of subspecies is that there exists no common understanding what constitutes this category, and some authors suggested therefore abandoning the usage of subspecies (e.g. Cracraft, 1983; Frost & Hillis, 1990; Isaac *et al*., 2004; van Nieukerken *et al*., 2016), with the result that subspecies are rarely used for invertebrates (but see Braby *et al*., 2012 for butterflies and Wallin *et al*., 2017 for beetles), fishes (Fricke *et al*., 2018), and amphibians (Frost, 2018). In a nutshell, the debate can be focused upon the question '*Is a subspecies merely a* (*local*) *variant without evolutionary significance, or does a subspecies represent a genetic evolutionarily significant lineage?*'

Several authors have reviewed aspects of the subspecies debate (Phillimore & Owens, 2006; Braby *et al*., 2012; Torstrom *et al*., 2014; Zachos, 2016), and the interested reader is referred to these publications. In the case of *Natrix helvetica*, abandoning the usage of subspecies would imply lumping together morphologically and genetically clearly distinct taxa that are allopatric or parapatric, including the morphologically highly distinc-

tive grass snakes from Corsica and Sardinia. It would also imply lumping together endangered lineages, in particular the imperiled grass snake from Sardinia (Vanni & Cimmaruta, 2011), and abundant lineages. This would be clearly counterproductive for conservation. Recently, Hawlitschek *et al*. (2012) discussed the merits of the rank of subspecies, especially with respect to allopatric island taxa that have reached a level of divergence "that does not yet warrant species status." We fully agree with their reasoning but see good arguments for the recognition of subspecies also beyond island taxa, especially in the light that subspecies typically originate as the result of range disjunctions, i.e. in allopatric situations resembling those on islands.

Our understanding of subspecies is similar to Moritz's (1994) definition of *Evolutionarily Significant Units* (*ESUs*), resembling the views of Braby *et al*. (2012) and Torstrom *et al*. (2014). Accordingly, subspecies should represent distinct mtDNA lineages (except for cases of mitochondrial capture), and they should represent also distinct nuclear genomic clusters, i.e. they should be confirmed by two independent lines of genetic evidence. Ideally, but not necessarily, distinct subspecies should be diagnosable also morphologically. In contrast to species, and in agreement with the definition for subspecies of Avise & Ball (1990), subspecies differ from species in that they are capable of extensive gene flow, i.e. they are completely genetically compatible with other subspecies. In our definition, subspecies constitute an early stage in the speciation process that has not reached yet a pronounced state of reproductive isolation. Thus, our subspecies concept is firmly embedded within the Biological Species Concept coined first by Mayr (1942), inspired by Rensch's (1929, 1947) *Rassenkreis* concept. In parapatric taxa, gene flow across contact zones represents the test case for applying the rank of a species or subspecies.

Populations representing distinct subspecies have gained genetic divergence in allopatry and have developed a distinct genetic profile, but are still fully capable of admixture. Therefore, subspecies can disappear as a consequence of genetic admixture during secondary contact. This process is similar to despeciation (Grant & Grant, 2014) or reverse speciation (Seehausen, 2006; Webb *et al*., 2011) – if the involved taxa were correctly classified as distinct species – but in subspecies, complete genetic amalgamation should be the rule and not the exception. The cases of despeciation and reverse speciation make also clear that the evolutionary transition between subspecies and species is gradual and that classification may fail. The advantage of our understanding of subspecies compared to ESUs is that the usage of subspecies facilitates communication in science and beyond. Often, recognizing subspecies continues also the usage of well-established scientific names, especially in some vertebrate groups, and helps to implement legislative measures for endangered subspecies in national and international law.

Grass snakes represent an excellent example for our subspecies concept because this group embraces taxa that represent different stages in the speciation process. *Natrix astreptophora* and *N. helvetica* are old evolutionary lineages and no longer capable of genetic admixture. According to a fossil-calibrated molecular clock, *N. astreptophora* diverged from the common ancestor of *N. natrix* and *N. helvetica* 10.61 mya (mean estimate; Kindler *et al*., 2018b). In the contact zone of *N. astreptophora* and *N. helvetica* hybrids occur only sporadically (Pokrant *et al*., 2016). *Natrix helvetica* and *N. natrix* are genetically less divergent and represent phylogenetically younger taxa (mean estimate of divergence: 8.64 mya; Kindler *et al*., 2018b). In their contact zone, hybridization occurs regularly, but the contact zone is bimodal, with parental genotypes still present, besides hybrid individuals. Moreover, the contact zone is narrow, only approximately 50 km wide, and gene flow is largely unidirectional from *N. helvetica* into *N. natrix* (Kindler *et al*., 2017). Within *N. natrix*, several distinct genetic lineages exist, among others the so-called 'red lineage' and 'yellow lineage', that represent mtDNA lineages matching with distinct nuclear genomic clusters (Kindler *et al*., 2017). The divergence of the red and yellow lineages is younger than that between *N. helvetica* and *N. natrix* (6.74 mya; Kindler *et al*., 2018b). Their contact zone is unimodal, with hybrid genotypes but without parental genotypes, and wide (approximately 680 km), with extensive wide-ranging gene flow in both directions (Kindler *et al*., 2017). Accordingly, the red and the yellow lineages qualify formally as subspecies, while the taxa without or restricted unidirectional gene flow can be recognized full species.

***Subspecies of* Natrix helvetica (*Lacepède, 1798*).** Only limited material is available for the southern lineages of *Natrix helvetica* and additional sampling may refine our conclusions. Not all mitochondrial lineages of *N. helvetica* match with nuclear genomic clusters (Fig. 8). This is indirect evidence for gene flow, reflected by lacking nuclear genomic differentiation. Mitochondrial DNA (Kindler *et al*., 2013) revealed lineage A (Sicily and Calabria) as sister group to the remaining mtDNA lineages. Thus, it should be assumed that grass snakes from Sicily and Calabria are also highly divergent with respect to nuclear genomic markers. However, microsatellite data indicated that Sicilian populations are less differentiated from mainland lineages than those from Corsica and Sardinia (Fig. 8), even though grass snakes from these two islands are less divergent in mtDNA. This conflicting pattern can be explained by past gene flow between Sicilian and peninsular Italian grass snakes across the narrow Strait of Messina. As a consequence of the much wider sea strait separating Corsica and Sardinia from mainland Italy, admixture between peninsular and Corso-Sardinian populations was more restricted, leading to more pronounced nuclear genomic divergence.

Our relatively few samples from mainland Italy often prevent direct evidence for gene flow among the mainland lineages. However, the following observations argue for their conspecificity: (1) There is clear evidence

for admixture between lineage E and lineage F in Northwestern Italy (Figs 3 and 4); (2) the microsatellite profiles of lineages C and F are not distinct (Figs 3 and 5); and (3) lineage A clusters in STRUCTURE analyses with the other lineages from mainland Italy. This suggests that all lineages, including lineage B from Corsica and Sardinia, should be regarded as conspecific because, if the most divergent and oldest mitochondrial lineage A is conspecific with the other mainland lineages, also the younger lineage B should be conspecific (cf. Fig. 8). This is also supported by STRUCTURE analyses for samples representing all mitochondrial lineages that place Corso-Sardinian samples in the same cluster than Sicilian and mainland Italian samples (Fig. 3, top).

Since GENÉ (1839), grass snakes from Corsica and Sardinia were always regarded as taxonomically distinct, and since MERTENS (1957), grass snakes from Corsica and Sardinia were recognized as two distinct subspecies (e.g. MERTENS & WERMUTH, 1960; THORPE, 1979; KABISCH, 1999; KREINER, 2007; VANNI & CIMMARUTA, 2011; GENIEZ, 2015). However, mitochondrial haplotypes from each island were not reciprocally monophyletic (KINDLER *et al*., 2013), even though we found no shared haplotypes for cyt *b* in the present study. Moreover, most of our microsatellite analyses could not discriminate between snakes from Corsica and Sardinia; only our PCA revealed distinct clusters (Figs 5 and 8), but the reliability of this result is restricted by the small sample size (Corsica: $n=3$; Sardinia: $n=2$). Therefore, Corsican and Sardinian grass snakes do not fully qualify for all of our criteria for the recognition of distinct subspecies. However, we are reluctant to lump together the two subspecies, also because Sardinian grass snakes are rare and endangered (VANNI & CIMMARUTA, 2011). Instead, we encourage further research and propose the continued tentative recognition of a distinct subspecies for each island.

Problematical is also the status of grass snakes from Northern Italy and Apulia. GenBank sequences from a single individual from Southern Apulia (GUICKING *et al*., 2006) indicate that a distinct mtDNA lineage (D) occurs there, which diverged more than 1 mya from its sister group (KINDLER *et al*., 2018b; Fig. 8). However, this lineage could not be studied using microsatellites. In Northern Italy and adjacent Switzerland, another mitochondrial lineage (C) was found (KINDLER *et al*., 2013; this study). In microsatellite analyses, lineage C could not be discriminated from lineage F which matches with the putative range of *N. h. lanzai* (KRAMER, 1970; Fig. 8). Morphologically, KRAMER (1970) assigned populations from the range of lineage C to what is now *N. h. helvetica*, a taxon regarded by him as distinct from his newly described subspecies *lanzai* from further south. Using multivariate analyses of many morphological characters, THORPE (1979) rejected the validity of *N. h. lanzai* and identified populations from the Padan Plain (Northern Italy) as representing a hybrid zone of what is now *N. helvetica* and *N. natrix*. This was not corroborated by genetics (KINDLER *et al*., 2013). However, for the Rhine region, there was a perfect match between the contact zone of *N. helvetica* and *N. natrix* as identified by THORPE (1979) and the contact zone revealed by genetics (KINDLER *et al*., 2017). Therefore, the conflict of morphological and genetic data for Northern Italy is unexpected and could be due to morphological peculiarities of the Northern Italian grass snakes. This calls for further research. For the time being, we tentatively assign grass snakes of lineage C to *N. h. lanzai* because our microsatellite analyses failed to discriminate it from this subspecies. Pending further research, we also recommend to identify Apulian grass snakes with *N. h. lanzai*. We are aware that this preliminary subspecies delineation renders the mitochondrial lineages of *N. h. lanzai* paraphyletic with respect to *N. h. helvetica*. However, this situation could change if further studies should reveal further lines of evidence supporting the taxonomic distinctness of Apulian and Northern Italian grass snakes. Until this issue is cleared, each population of *N. h. lanzai* harboring a distinct mitochondrial lineage should be regarded as a Management Unit. In conclusion, we suggest recognizing tentatively the following five subspecies of *Natrix helvetica* (Lacepède, 1789):

1) *Natrix helvetica helvetica* (Lacepède, 1789), Western Europe from the Pyrenees to the Rhine region, including Britain.
2) *Natrix helvetica cetti* Gené, 1839, Sardinia.
3) *Natrix hevetica corsa* (Hecht, 1930), Corsica.
4) *Natrix helvetica lanzai* Kramer, 1970, Po drainage and peninsular Italy, except for parts of Calabria.
5) *Natrix helvetica sicula* (Cuvier, 1829), Calabria (in part), Sicily.

The taxonomic identity of grass snakes from Northern Italy, Apulia, Corsica, and Sardinia requires further study. The records of the mitochondrial lineage C of *N. h. lanzai* in the Swiss canton Valais (Wallis), beyond the Po drainage, and in considerable altitude in the canton Ticino (1,900 m a.s.l.) suggest that this lineage might have crossed the main chain of the Alps also elsewhere. An unstudied contact zone of *N. h. lanzai* and *N. natrix* is expected for northeasternmost Italy, from where records of *N. natrix* are known east of the Piave River (LAPINI *et al*., 1999).

Acknowledgements

This study was funded by the German Research Foundation (DFG; FR 1435/11-1 and FR 1435/11-2). Many thanks go to all colleagues who donated samples from the southern distribution range of *Natrix helvetica*: Markus Auer, Birgit Blosat, Maxime Chèvre, Marc Cheylan, Marco Favelli, Michael Franzen, Philippe Geniez, Frank Glaw, Jakob Hallermann, Axel Hille, Daniel Jablonski, Ulrich Scheidt, and Sylvain Ursenbacher. Wolfgang Böhme and Philippe Geniez provided the photos of grass snakes shown in our figures. Daniel Jablonski and Edoardo Razzetti reviewed the manuscript.

References

Avise, J. C. & Ball, R. M. (1990). Principles of genealogical concordance in species concepts and biological taxonomy, pp. 45–67 *in*: Futuyma, D. & Antonovics, J. (eds) *Oxford Surveys in Evolutionary Biology*. Oxford, UK, Oxford University Press.

Bannikov, A. G., Darevsky, I. S., Ishchenko, V. G., Rustamov, A. K. & Shcherbak, N. N. (1977). *Opredelitel' zemnovodnykh i presmykayushchikhsya fauny SSSR*. Moscow, Prosveshcheniye, 414 pp.

Braby, M. F., Eastwood, R. & Murray, N. (2012). The subspecies concept in butterflies: has its application in taxonomy and conservation biology outlived its usefulness? *Biological Journal of the Linnean Society*, **106**, 699–716.

Braña, F. (1998). *Natrix natrix* (Linnaeus, 1758), pp. 454–466 *in*: Salvador, A. (ed.) *Fauna ibérica*, *Vol. 10*. *Reptiles*. Madrid, Museo Nacional de Ciencias Naturales and Consejo Superior de Investigaciones Científicas.

Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Clements, J. F., Schulenberg, T. S., Iliff, M. J., Roberson, D., Fredericks, T. A., Sullivan, B. L. & Wood, C. L. (2018). *The eBird/Clements Checklist of Birds of the World: v2018*. http://www.birds.cornell.edu/clementschecklist/download/, accessed 14 September 2018.

Cracraft, J. (1983). Species concepts and speciation analysis, pp. 159–187 *in*: Johnston, R. F. (ed.) *Current Ornithology*, *Vol 1*. New York, Plenum Press.

Earl, D. A. & vonHoldt, B. M. (2012). structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Engelmann, W.-E., Fritzsche, J., Günther, R. & Obst, F. J. (1986). *Lurche und Kriechtiere Europas*. Stuttgart, Ferdinand Enke Verlag, 420 pp.

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.

Excoffier, L. & Lischer, H. E. L. (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.

Falush, D., Stephens, M. & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

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

Fricke, R., Eschmeyer, W. N. & van der Laan, R. (2018, eds). *Catalog of Fishes: Genera*, *Species*, *References*. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp, accessed 14 September 2018.

Frost, D. R. (2018). *Amphibian Species of the World: an Online Reference*. *Version 6.0*. http://research.amnh.org/herpetology/amphibia/index.html, accessed 14 September 2018.

Frost, D. R. & Hillis, D. M. (1990). Species concepts and practice: herpetological applications. *Herpetologica*, **46**, 87–104.

Gené, J. (1839). Synopsis Reptilium Sardiniae indigenorum. *Memorie della Reale Accademia delle Scienze di Torino*, **1** (1838), 257–286 + 5 pls.

Geniez, P. (2015). *Serpents d'Europe*, *d'Afrique du Nord et du Moyen-Orient*. Paris, Delachaux et Niestlé, 380 pp.

Grant, P. R. & Grant, B. R. (2014). *40 Years of Evolution*. *Darwin's Finches on Daphne Major*. Princeton, NJ, Princeton University Press, xxxii + 400 pp.

Gruber, U. (1989). *Die Schlangen Europas und rund ums Mittelmeer*. *Kosmos Naturführer*. Stuttgart, Franckh'sche Verlagshandlung, W. Keller & Co., 248 pp.

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.

Guicking, D., Joger, U. & Wink, M. (2008). Molekulare Phylogenie und Evolutionsgeschichte der Gattung *Natrix*, mit Bemerkungen zur innerartlichen Gliederung von *N. natrix*. *Mertensiella*, **17**, 16–30.

Hall, T. A. (1999). bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hawlitschek, O., Nagy, Z. T. & Glaw, F. (2012). Island evolution and systematic revision of Comoran snakes: why and when subspecies still make sense. *PLoS ONE*, **7**, e42970.

ICZN [International Commission on Zoological Nomenclature] (1999). *International Code of Zoological Nomenclature*. *Fourth Edition*. London, International Trust for Zoological Nomenclature, XXIX + 306 pp.

Isaac, N. J. B., Mallet, J. & Mace, G. M. (2004). Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology & Evolution*, **19**, 464–469.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Kabisch, K. (1999). *Natrix natrix* (Linnaeus, 1758) – Ringelnatter, pp. 513–580 *in*: Böhme, W. (ed.) *Handbuch der Reptilien und Amphibien Europas*. *Band 3/IIA: Schlangen II*. Wiebelsheim, Aula-Verlag.

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. megalocephala*). *Zoologica Scripta*, **42**, 458–472.

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.

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.

Kindler, C., Graciá, E. & Fritz, U. (2018a). Extra-Mediterranean glacial refuges in barred and common grass snakes (*Natrix helvetica*, *N. natrix*). *Scientific Reports*, **8**, 1821.

Kindler, C., de Pous, P. Carranza, S., Beddek, M., Geniez, P. & Fritz, U. (2018b). Phylogeography of the Ibero-Maghrebian red-eyed grass snake (*Natrix astreptophora*). *Organisms*, *Diversity & Evolution*, **18**, 143–150.

Kramer, E. (1970). Revalidierte und neue Rassen der europäischen Schlangenfauna. *Lavori della Società Italiana di Biogeografia*, **1**, 667–676, pls. I–VIII.

Kramer, E., Linder, A. & Mermillod, B. (1982). Systematische Fragen zur europäischen Schlangenfauna. *Vertebrata Hungarica*, **21**, 195–201.

Kreiner, G. (2007). *Schlangen Europas*. Frankfurt am Main, Edition Chimaira, 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.

Lapini, L., dall'Asta, A., Bressi, N., Dolce, S. & Pellarini, P. (1999). *Atlante corologico degli Anfibi e dei Rettili del Friuli-Venezia Giulia*. Udine, Edizioni del Museo Friulano di Storia Naturale, 149 pp.

Mayr, E. (1942). *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. New York, Columbia University Press, xii + 334 pp.

Mertens, R. (1947). Studien zur Eidonomie und Taxonomie der Ringelnatter (*Natrix natrix*). *Abhandlungen der Senckenbergischen Naturforschenden Gesellschaft*, **476**, 1–38.

Mertens, R. (1957). Die Amphibien und Reptilien Korsikas. *Senckenbergiana biologica*, **38**, 175–192.

Mertens, R. & Wermuth, H. (1960). *Die Amphibien und Reptilien Europas* (*Dritte Liste*, *nach dem Stand vom 1. Januar 1960*). Frankfurt am Main, Verlag Waldemar Kramer, XI + 264 pp.

Moritz, C. (1994). Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution*, **9**, 373–375.

van Nieukerken, E. J., Doorenweerd, C., Hoare, R. J. B. & Davis, D. R. (2016). Revised classification and catalogue of global Nepticulidae and Opostegidae (Lepidoptera, Nepticuloidea). *ZooKeys*, **628**, 65–246.

van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. (2004). micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Orlov, N. L. & Tuniyev, B. S. (1987). Novyi vid uzha *Natrix megalocephala* sp. nov. s Kavkaza (Ophidia: Colubridae). *Trudy Zoologicheskogo Instituta Akademii Nauk SSSR*, **158**, 116–130.

Phillimore, A. B. & Owens, I. P. F. (2006). Are subspecies useful in evolutionary and conservation biology? *Proceedings of the Royal Society B*, **273**, 1049–1053.

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.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Puechmaille, S. J. (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.

Rensch, B. (1929). *Das Prinzip geographischer Rassenkreise und das Problem der Artbildung*. Berlin, Gebrüder Borntraeger, 206 pp.

Rensch, B. (1947). *Neuere Probleme der Abstammungslehre*. Stuttgart, Ferdinand Enke Verlag, VI + 407 pp.

Rosenberg, N. A. (2004). distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Seehausen, O. (2006). Conservation: losing biodiversity by reverse speciation. *Current Biology*, **16**, R334–R337.

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*. Latina, Edizioni Belvedere, 543 pp.

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

Thorpe, R. S. (1979). Multivariate analysis of the population systematics of the ringed snake, *Natrix natrix* (L.). *Proceedings of the Royal Society of Edinburgh*, **78B**, 1–62.

Thorpe, R. S. (1980). Microevolution and taxonomy of European reptiles with particular reference to the grass snake *Natrix natrix* and the wall lizards *Podarcis sicula* and *P. melisellensis*. *Biological Journal of the Linnean Society*, **14**, 215–233.

Thorpe, R. S. (1984a). Primary and secondary transition zones in speciation and population differentiation: a phylogenetic analysis of range expansion. *Evolution*, **38**, 233–243.

Thorpe, R. S. (1984b). Geographic variation in the western grass snake (*Natrix natrix helvetica*) in relation to hypothesized phylogeny and conventional subspecies. *Journal of Zoology*, **203**, 345–355.

Torstrom, S. M., Pangle, K. L. & Swanson, B. J. (2014). Shedding subspecies: the influence of genetics on reptile subspecies taxonomy. *Molecular Phylogenetics and Evolution*, **76**, 134–143.

Uetz, P., Freed, P. & Hošek, J. (2018, eds). *The Reptile Database*. http://www.reptile-database.org, accessed 14 September 2018.

Vanni, S. & Cimmaruta, R. (2011). *Natrix cetti* Gené, 1839, pp. 538–545 *in*: Corti, C., Capula, M., Luiselli, L., Razzetti, E. & Sindaco, R. (eds), *Fauna d'Italia*, *Reptilia*. Bologna, Calderini.

Wallin, H., Kvamme, T. & Bergsten, J. (2017). To be or not to be a subspecies: description of *Saperda populnea lapponica* ssp. n. (Coleoptera, Cerambycidae) developing in downy willow (*Salix lapponum* L.). *ZooKeys*, **691**, 101–148.

Webb, W. C., Marzluff, J. M. & Omland, K. E. (2011). Random interbreeding between cryptic lineages of the common raven: evidence for speciation in reverse. *Molecular Ecology*, **20**, 2390–2402.

Wilson, D. E. & Reeder, D. M. (2005, eds). *Mammal Species of the World*. *A Taxonomic and Geographic Reference* (*3rd ed.*). Baltimore, MD, Johns Hopkins University Press, 2,142 pp.

Zachos, F. E. (2016). *Species Concepts in Biology*. *Historical Development*, *Theoretical Foundations and Practical Relevance*. Basel, Springer, xii + 220 pp.

Electronic Supplement Files

at http://www.senckenberg.de/vertebrate-zoology

File 1. Natrix_helvetica_Supporting_Information.pdf.

File 2. Natrix_helvetica_Supporting_Table_S1.xlsx.