Org Divers Evol (2012) 12:71–80
DOI 10.1007/s13127-011-0069-8

ORIGINAL ARTICLE

ORGANISMS DIVERSITY & EVOLUTION

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

Uwe Fritz · Claudia Corti · Martin Päckert

Received: 29 July 2011 / Accepted: 28 November 2011 / Published online: 8 January 2012

Abstract We supplement a previously published mitochondrial DNA data set of grass snake sequences (ND1, ND2, ND4, cyt *b*, in total 3,806 bp) with sequences of Corso-Sardinian and Tuscan specimens and infer their phylogeny using Bayesian, maximum likelihood and maximum parsimony methods. In addition, we estimate divergence times of grass snake clades using a relaxed molecular clock calibrated with fossil evidence, and, in a second approach, the post-Messinian reopening of the Strait of Gibraltar. Recently it was suggested that Corso-Sardinian grass snakes represent a distinct species: *Natrix cetti*. All tree-building methods revealed well-supported branching patterns and deep divergences among grass snakes. However, sequences of *N. natrix* were consistently paraphyletic with respect to Corso-Sardinian sequences. The sister group of Corso-Sardinian grass snakes is a clade embracing *N. n. helvetica* and *N. n. lanzai*. Extensive gene flow between *N. n. helvetica* and a more distantly related subspecies (*N. n. natrix*) is well known, which is why we conclude that the status of Corso-Sardinian grass snakes as subspecies of *N. natrix* should be reinstated. Many currently recognized grass snake subspecies conflict with mitochondrial clades, suggestive of inappropriate morphological taxon delineation and mitochondrial introgression. Divergences among grass snakes are old, and the results of the two independent dating approaches are largely congruent. Accordingly, the Alpine orogenesis seems to have caused the origin of the oldest clade, corresponding to Iberian *N. n. astreptophora*. The formation of Corso-Sardinian grass snakes was dated to the Early Pliocene and could result from post-Messinian flooding of the Mediterranean Basin. Another deeply divergent clade of approximately the same age, endemic in central and northern Europe, suggests the Pleistocene survival of grass snakes north of the Alps. At least one glacial refuge in which old lineages survived Pleistocene cold periods was located on each of the three major southern European peninsulas and in Anatolia. Due to pronounced sequence divergences among Italian and southern Swiss grass snakes, we hypothesize multiple refugia south of the Alps and in the Apennine Peninsula, and there is evidence for two refuges on the Balkan Peninsula.

Keywords Reptilia · Squamata · Serpentes · *Natrix natrix*

U. Fritz (✉) · M. Päckert
Museum of Zoology (Museum für Tierkunde),
Senckenberg Dresden,
A. B. Meyer Building,
01109 Dresden, Germany
e-mail: uwe.fritz@senckenberg.de

C. Corti
Museo di Storia Naturale dell'Università di Firenze,
Sezione di Zoologia "La Specola",
Via Romana, 17,
50125 Florence, Italy

Introduction

The grass snake, *Natrix natrix* (Linnaeus 1758), is a widely distributed colubrid species, ranging from North Africa and the Iberian Peninsula over most of Europe to Lake Baikal in Central Asia. Grass snakes are associated closely with water, and swim well. They feed mainly on amphibians and fishes, but take also small mammals and nestling birds (Arnold and Burton 1978; Kabisch 1999). Based on morphological characters, many subspecies have been described, and some recent authors have recognised up to 14 distinct subspecies (Blosat 2008; Gruber 1989; Kabisch 1999; Kreiner 2007). In contrast, based on multivariate analyses of phenotypic

characters, Thorpe (1979) accepted only four subspecies. According to Thorpe (1979), the subspecies *N. n. helvetica* occupies the western part of the distribution range, from North Africa and the Iberian Peninsula to Central Europe, and *N. n. natrix* the eastern part. Both subspecies meet and hybridize in the Rhine region. Besides *N. n. natrix* and *N. n. helvetica*, Thorpe (1979) recognized the highly distinctive grass snakes from Sardinia and Corsica as the subspecies *N. n. cetti* and *N. n. corsa*, respectively. Using allozyme data, Hille (1997) confirmed the hybridization of *N. n. natrix* and *N. n. helvetica* in the Rhine region and the distinctiveness of Sardinian and Cypriote grass snakes. Later, Guicking (2004) and Guicking et al. (2006, 2008) published a comprehensive data set of mtDNA sequences that was used to study phylogeny and divergence times within the genus *Natrix* and for a preliminary assessment of the subspecies and phylogeography of *N. natrix*. Due to the rareness of Sardinian and Corsican grass snakes, Guicking (2004) and Guicking et al. (2006, 2008) were not able to include these taxa. Guicking et al. (2008) concluded that Thorpe's (1979) four subspecies model underestimates the actual taxonomic differentiation, but that further sampling is needed before drawing definite taxonomic conclusions. Acknowledging the pronounced morphological distinctiveness of Corso-Sardinian grass snakes, Vanni and Cimmaruta (2010) recently transferred these taxa to the species *N. cetti* (with the subspecies *N. c. cetti* and *N. c. corsa*). Thus, the two competing schemes of subspecies delineation of *N. natrix* represent an unresolved problem that was further aggravated by the separation of *N. cetti*.

In the present paper we follow the traditional subspecies delineation of *N. natrix* and recognise among the studied grass snakes the subspecies listed in Table 1. However, acknowledging the latest revision by Vanni and Cimmaruta (2010), we provisionally treat the Corso-Sardinian grass snakes as representing the distinct species *N. cetti* with the two subspecies *N. c. cetti* and *N. c. corsa*. To assess their taxonomic status, we supplement the data set of Guicking (2004) and Guicking et al. (2006) with homologous mitochondrial DNA sequences of one Tuscan, one Sardinian, and two Corsican grass snakes (ND1, ND2, ND4, cyt *b*, in total 3,806 bp) and use this expanded data set for phylogenetic analyses. Even though none of the previous investigations on grass snakes explicitly used a clearly defined species concept, the Biological Species Concept (cf. Coyne and Orr 2004; Mayr 1942)—the prevailing species concept in European herpetology—was implicitly applied. Within that framework, distinct evolutionary entities are treated as subspecies when there is extensive gene flow among them, as suggested by morphological and genetic evidence for *N. n. natrix*, *N. n. helvetica* and *N. n. persa* (Hille 1997; Kabisch 1999; Kreiner 2007; Thorpe 1979). In the absence of direct evidence (as in allopatric taxa like the island snakes *N. c. cetti* and *N. c. corsa*), gene flow among parapatric subspecies may be used as a yardstick. Consequently, if *N. cetti* represents a distinct species, a significantly greater genetic divergence were expected in comparison to subspecies with evident gene flow. In phylogenetic analyses, this should correspond to a sister group relationship between *N. cetti* and *N. natrix*.

In addition to the taxonomic reappraisal of *N. cetti*, we use the expanded data set of Guicking (2004) and Guicking et al. (2006) to re-evaluate the phylogeography of grass snakes. We apply a relaxed molecular clock to date the splits between mitochondrial lineages, in particular to test the hypothesis that Corso-Sardinian grass snakes are of ancient, perhaps pre-Pleistocene origin (Lanza 1988). For calibrating our clock, we use a recently discovered Sardinian *Natrix* fossil (Delfino et al. 2011) and, in an independent calculation, the post-Messinian reopening of the Strait of Gibraltar (García-Castellanos et al. 2009)—an event often used for calibrating molecular phylogenies.

Materials and methods

Laboratory procedures

We extracted total genomic DNA from tissue samples using the DTAB method (Gustincich et al. 1991) and amplified and sequenced the same four mtDNA fragments as in Guicking et al. (2006), corresponding to the genes coding for the NADH dehydrogenase subunits 1, 2, and 4 (ND1, ND2, ND4) and to the cytochrome *b* gene (cyt *b*). For PCR and sequencing, we applied the primer pairs 16Sb+tRNA-ile (ND1), L4437b+tRNA-trp(ND2), ND4ab+tRNA-leu (ND4) and L14724NAT+Thrsnr2 (cyt *b*) of Guicking et al. (2006).

We amplified each gene in a total volume of 20 μl containing 1 unit *Taq* polymerase (Bioron, Ludwigshafen, Germany), 1 x buffer (as recommended by the supplier), 0.5 μM of each primer, and 0.2 mM of each dNTP (Fermentas, St. Leon-Rot, Germany) using the same PCR conditions, with initial denaturation at 94°C for 5 min, followed by 35 cycles with denaturation at 94°C for 45 s, annealing at 55°C for 45 s, and extension at 72°C for 60 s. After the last cycle, we incubated samples in a final extending step at 72°C for 10 min. We purified PCR products using the ExoSAP-IT enzymatic cleanup (USB Europe, Staufen, Germany; modified protocol: 30 min at 37°C, 15 min at 80°C) and sequenced them on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA) using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

The lengths of the resulting mtDNA sequences were 980–1,083 bp (ND1), 949–975 bp (ND2), 829–883 bp (ND4), and 1,070–1,101 bp (cyt *b*).

Table 1 GenBank sequences and samples used in the present study. Subspecies allocation according to Kabisch (1999). Museum acronyms for vouchers follow Sabaj Pérez (2010)

Species/subspecies	Collecting site	Voucher	GenBank accession numbers				Reference
			ND1	ND2	ND4	cyt *b*	
Natrix cetti cetti Gené 1839	Italy: Sardinia: Belvì (Nuoro)	MZUF 40268	HE584617	HE584618	HE584619	HE584620	This study
Natrix cetti corsa (Hecht 1930)	France: Corsica: Porto-Vecchio	MTD 35388	HE584621	HE584622	HE584623	HE584624	This study
Natrix cetti corsa (Hecht 1930)	France: Corsica: Santa Giulia	MTD 42489	HE584625	HE584626	HE584627	HE584628	This study
Natrix natrix astreptophora (Seoane 1884)	Spain: Cádiz Province	MNCN 13796	AY873748	AY870621	AY873714	AY866535	Guicking et al. (2006)
Natrix natrix astreptophora (Seoane 1884)	Spain: Catalonia: Osona	MVZ 200534	AY873747	AY870620	AY873713	AY866536	Guicking et al. (2006)
Natrix natrix helvetica (Lacepède 1789)	England: Kent: Isle of Sheppey	LSUMZ 41506	AY873743	AY870617	AY873710	AY866544	Guicking et al. (2006)
Natrix natrix helvetica (Lacepède 1789)	France: vicinity of Paris	———————	AY873760	AY870640	AY873736	AY866537	Guicking et al. (2006)
Natrix natrix helvetica (Lacepède 1789)	Switzerland: Ticino: Astano	———————	AY487760	AY487780	AY487795	AY487751	Guicking et al. (2006)
Natrix natrix helvetica x *natrix*	Germany: Kelkheim/Taunus	———————	AY873744	AY870618	AY873711	AY866538	Guicking et al. (2006)
Natrix natrix helvetica x *natrix*	Germany: Lake Constance: Radolfzell	———————	AY873745	AY487777	AY487792	AY487727	Guicking et al. (2006)
Natrix natrix lanzai Kramer 1970	Italy: Apulia: Torre San Gennaro (Brindisi)	———————	AY873749	AY870622	AY873715	AY487733	Guicking et al. (2006)
Natrix natrix lanzai Kramer 1970	Italy: Tuscany: Vaglia (Firenze)	MZUF 40267	HE584629	HE584630	HE584631	HE584632	This study
Natrix natrix natrix (Linnaeus 1758)	Denmark: Jutland	———————	AY873746	AY870619	AY873712	AY866539	Guicking et al. (2006)
Natrix natrix natrix (Linnaeus 1758)	Russia: Kaliningrad Oblast: Rybachy	———————	AY873758	AY487778	AY487793	AY487741	Guicking et al. (2006)
Natrix natrix natrix (Linnaeus 1758)	Sweden: Småland: Högsby	———————	AY487764	AY487784	AY487799	AY487755	Guicking et al. (2006)
Natrix natrix natrix x *persa*	Hungary: Pécsi-tó	———————	AY487761	AY487781	AY487796	AY487752	Guicking (2004)
Natrix natrix natrix x *persa*	Romania: Cluj-Napoca	———————	AY873753	AY870625	AY873719	AY866541	Guicking et al. (2006)
Natrix natrix natrix x *persa*	Romania: Tulcea Region	———————	AY873752	AY870624	AY873718	AY866540	Guicking et al. (2006)
Natrix natrix natrix x *persa*	Slovenia: Zalec	———————	AY873754	AY870626	AY873720	AY487738	Guicking et al. (2006)
Natrix natrix persa (Pallas 1814)	Armenia: Ankavan	ROM 26842	AY873756	AY870629	AY873723	AY866543	Guicking et al. (2006)
Natrix natrix persa (Pallas 1814)	Bulgaria: Malko Tarnovo District	CAS 219930	AY873751	AY870623	AY873717	AY866542	Guicking et al. (2006)
Natrix natrix persa (Pallas 1814)	Georgia: Batumi	———————	AY873759	AY870631	AY873725	AY487736	Guicking et al. (2006)
Natrix natrix persa (Pallas 1814)	Greece: Ioánnina	———————	AY873750	———————	AY873716	AY487725	Guicking et al. (2006)
Natrix natrix persa (Pallas 1814)	Iran: Kermanshah Province	———————	———————	———————	AY487800	AY487756	Guicking (2004)
Natrix natrix persa (Pallas 1814)	Turkey: Hattuşa (Sarikale)	———————	AY873755	AY870628	AY873722	AY487726	Guicking et al. (2006)
Natrix natrix persa (Pallas 1814)	Turkey: Yeniçaga	———————	AY873737	AY870627	AY873721	AY487730	Guicking et al. (2006)
Natrix natrix scutata (Pallas 1771)	Kazakhstan: Emba River	———————	AY487759	AY487779	AY487794	AY487749	Guicking et al. (2006)
Natrix natrix scutata (Pallas 1771)	Russia: Penza District	———————	AY487763	AY487783	AY487798	AY487754	Guicking (2004)
Natrix natrix scutata (Pallas 1771)	Russia: Samara District	———————	AY487762	AY487782	AY487797	AY487753	Guicking et al. (2006)
Natrix natrix scutata (Pallas 1771)	Russia: Tula District	CAS 175878	AY873757	AY870630	AY873724	AF471059	Lawson et al. (2005), Guicking et al. (2006)
Natrix maura (Linnaeus 1758)	Spain: Southern Spain	MNCN 12016	AY873741	AY870615	AY873708	AY866530	Guicking et al. (2006)
Natrix maura (Linnaeus 1758)	Morocco: Tétouan Province	MVZ 178093	AY873742	AY870616	AY873709	AF420077	de Queiroz et al. (2002), Guicking et al. (2006)
Natrix tessellata (Laurenti 1768)	Armenia: Geolazar	ROM 23418	AY873769	AY870639	AY873734	AY866531	Guicking et al. (2006)
Nerodia fasciata (Linnaeus 1766)	USA: Florida: Citrus County	CAS 211010	AY873738	AY870612	AY873705	AY866529	Guicking et al. (2006)

Alignment, partitioning and phylogenetic analyses

We downloaded homologous GenBank sequences of known-locality grass snakes (Table 1), corresponding to the data sets of Guicking (2004) and Guicking et al. (2006), and aligned these with our newly generated sequences. As outgroups, we included sequences of *Natrix maura*, *N. tessellata*, and of the distantly related North American banded water snake (*Nerodia fasciata*). Acknowledging that mtDNA represents one and the same locus, we concatenated the four mtDNA fragments, resulting in an alignment of 3,806 bp total length. Position 1–1,117 of this alignment corresponded to cyt *b*; position 1,118–2,081, to ND1; position 2,082–3,110, to ND2; and position 3,111–3,806, to ND4. We determined the best evolutionary model for each partition in MrMODELTEST 2.3 (Nylander 2004) using the AIC, resulting in the GTR+G+I model for all partitions except ND4 for which GTR+G was suggested.

Using PAUP 4.0b10 (Swofford 2002), we reconstructed the phylogeny of *Natrix* sequences under maximum parsimony (MP) in a branch-and-bound search with gaps coded as fifth character state. We obtained clade support using the same search method and 1,000 bootstrap replicates (Felsenstein 1985), with all characters unordered and equally weighted; gaps were treated again as fifth character state.

Using RAxML 7.2.8 and the graphical user interface raxmlGUI 0.95 (Silvestro and Michalak 2011; Stamatakis 2006), we conducted maximum likelihood (ML) analyses, with data partitioned by gene. Applying the GTR+G+I model across all four partitions, we obtained for the best-scoring ML tree thorough bootstrap support using 1,000 replicates.

In addition, we performed Bayesian Analyses (BA) using MrBAYES 3.1.2 (Ronquist and Huelsenbeck 2003) and the implemented Metropolis-coupled Markov chain Monte Carlo algorithm with two parallel runs, each with one cold and three heated chains. We set the heating parameter λ to 0.1 and let the chains run for 10^6 generations, with every 100th generation sampled, allowing the two runs to converge onto the stationary distribution. Convergence was confirmed by average standard deviations of split frequencies approaching zero. We used a burn-in of 3,000 to sample only the plateau of the most likely trees for generating a 50% majority rule consensus tree. As in the ML analysis, we partitioned the data by gene and applied best-fit models to each of the four partitions. We allowed the overall rate to vary among partitions and set model parameters unlinked across partitions, so that for each partition a separate set of parameters was estimated.

Moreover, exploratory parsimony network analyses were calculated using TCS 1.21 (Clement et al. 2000) for subsets of the complete alignment. The final data matrix and all inferred phylogenetic trees are available from TreeBASE (http://www.treebase.org) under the ID 12048.

Molecular clock

We performed two independent molecular dating approaches using BEAST 1.4.8 (Drummond and Rambaut 2007) to estimate node ages and 95% highest posterior density (HPD) intervals. In the first approach, we used the end of the Messinian Salinity Crisis for dating the split between North African and European *Natrix maura* and applied a fixed minimum age of 5.3 million years (Ma; reopening of the Strait of Gibraltar) and a soft maximum of 5.96 Ma (onset of the Messinian Crisis) to the respective node. In the second calculation, we used a fossil age constraint for the split between Corso-Sardinian grass snakes and their closest continental relatives, corresponding to a Pliocene *Natrix* vertebra excavated in Sardinia (Delfino et al. 2011). We used its age (3.6 Ma) as minimum TMRCA prior constraint and the beginning of the Pliocene (5.3 Ma) as soft maximum. We did not combine the two calibration points in one analysis to assess whether the estimates obtained from either approach are similar. By doing so, we examine also the reliability of the reopening of the Strait of Gibraltar as a calibration point for water snakes of the genus *Natrix*.

In the two calculations using an uncorrelated relaxed clock (Drummond et al. 2006), we chose a lognormal prior distribution to capture the soft maximum constraint (Benton et al. 2009); lognormal means and standard deviations were accordingly adjusted and the minimum ages (fossil and palaeogeographic constraint, respectively) applied as zero offset values in order to match the distribution in real space. We set the MCMC chain to 10 million generations and let log parameters be sampled every 1,000th generation. We set the tree prior to speciation (yule process) and activated the "auto optimise" option to adjust automatically tuning parameters. We partitioned input sequence data manually in the XML file generated with BEAUTi according to the estimates of MrMODELTEST. We obtained linearised consensus trees including posterior probabilities using TREEANNOTATOR (as implemented in the BEAST package; Rambaut and Drummond 2007) with the burn-in parameter set to 3,000 and node heights to "mean heights".

Results

Phylogeny

The three applied tree-building methods revealed largely congruent topologies with high support values for most nodes (Fig. 1). The branching patterns of the ML and BA trees were in perfect agreement. Under MP, three equally

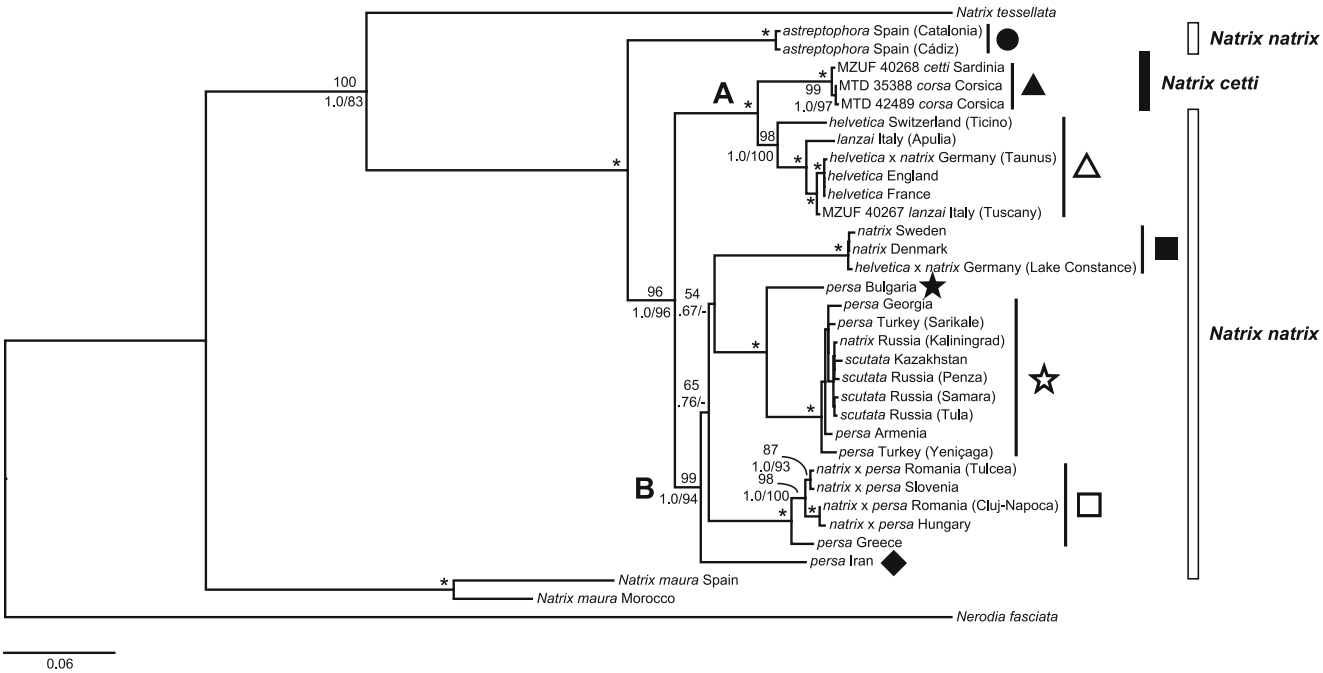

Fig. 1 Maximum likelihood (ML) tree for *Natrix* sequences calculated with RAxML based on 3,806 bp of mtDNA (ND1, ND2, ND4, cyt *b*). Numbers above nodes are thorough bootstrap values (RAxML); below nodes, Bayesian posterior probabilities and bootstrap values obtained under maximum parsimony (MP; not shown for some terminal clades with short branch lengths). For new samples, voucher codes (Table 1) are preceding taxon names. *Asterisks* indicate maximum support under all methods; *dashes*, branch not found. Symbols correspond to Fig. 4. Note the paraphyly of *N. n. helvetica* with respect to *N. n. lanzai* and the mismatches between subspecies and mitochondrial clades in clade B. On the right, species delineation of *N. cetti* and *N. natrix* according to Vanni and Cimmaruta (2010)

parsimonious trees were found (tree length: 2,281; CI= 0.679, RI=0.820) that differed only with respect to the sister group relations of the weakly supported branches of the other two analyses. Consequently, the strict consensus tree placed these branches in an unresolved polytomy.

The obtained phylogeny (Fig. 1) is in general agreement with the trees published by Guicking (2004) and Guicking et al. (2006, 2008) and contradicts the currently debated subspecies or species delineations within the *Natrix natrix* group. In particular, the most basal branch is not constituted by Corso-Sardinian grass snakes (*N. cetti*), as expected, but by the Iberian grass snake (*N. n. astreptophora*). The latter taxon is sister to a major clade containing sequences of all other grass snakes. This major clade consists of two other well-supported clades (A and B).

Clade A comprises the sequences of *N. cetti* that constitute a well-supported clade, albeit with very short terminal branch lengths. The sequences of the two Corsican grass snakes (*N. c. corsa*) are the weakly differentiated sister clade of the only Sardinian specimen (*N. c. cetti*). The Corso-Sardinian *Natrix* are with high support sister to a deeply structured, well-supported clade embracing sequences of *N. n. helvetica*, *N. n. lanzai* and a German snake from the hybrid zone between *N. n. helvetica* and *N. n. natrix*. Sequences of *N. n. helvetica* are paraphyletic with respect to *N. n. lanzai*. In parsimony network analysis (Fig. 2), sequences of *N. c. cetti* and *N. c. corsa* differ by a maximum of eight mutation steps, while among the sequences of the paraphyletic *helvetica-lanzai* clade a maximum of 96 steps occurs.

Clade B (Fig. 1) consists of several well-supported subclades whose basal branching patterns are poorly resolved, however. One subclade contains sequences of *N. n. natrix* from Denmark and Sweden plus the sequence of another *N. n. helvetica* x *natrix* hybrid from Germany. Another subclade comprises sequences of several subspecies, viz. *N. n. persa* from Armenia, Georgia, and Turkey, *N. n. natrix* from westernmost Russia (Kaliningrad Oblast) and *N. n. scutata* from several sites in Russia and Kazakhstan. Its well-supported sister is a highly divergent sequence of a *N. n. persa* from Bulgaria, representing a distinct terminal. A further distinct subclade corresponds to grass snakes from the hybrid zone between *N. n. natrix* and *N. n. persa* in Hungary, Romania and Slovenia and a *N. n. persa* from Greece. Another sequence of a *N. n. persa* from Iran is deeply divergent and excluded from any other subclade, so that clade B contains five distinct clades or deeply divergent terminals.

Molecular clock

The two calculations, based on two independent calibrations (post-Messinian reopening of the Strait of Gibraltar or fossil

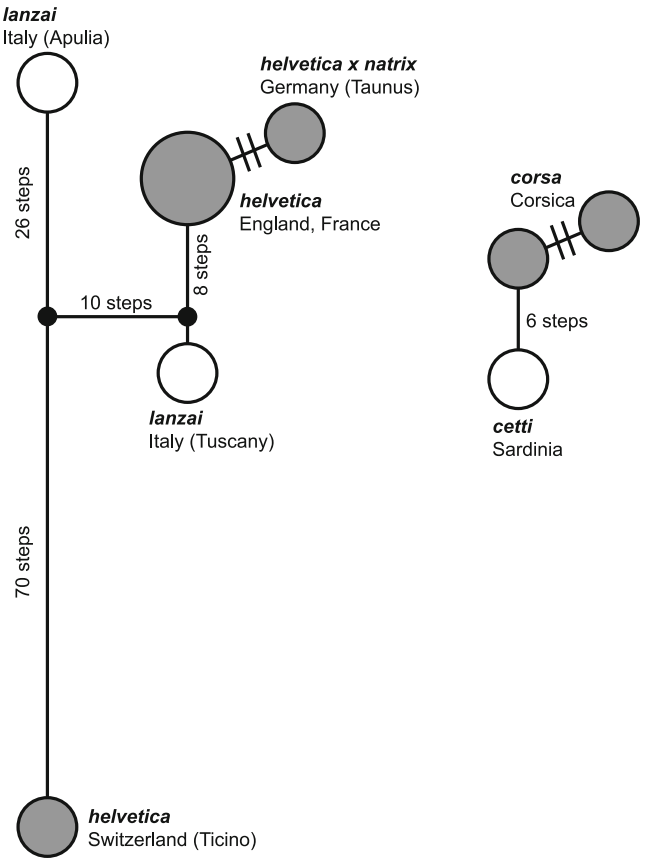

Fig. 2 Parsimony networks for haplotypes of *Natrix natrix helvetica*, *N. n. lanzai*, a *N. n. helvetica* x *natrix* hybrid (*left*) and Corso-Sardinian grass snakes (*right*) based on 3,806 bp of mtDNA (ND1, ND2, ND4, cyt *b*). The *large symbol* for *N. n. helvetica* indicates that this haplotype was found twice; *small black circles*, missing node haplotypes. Connections between haplotypes show number of mutation steps. Connection of haplotypes in left network enforced; 95% connection limit: 27 steps

record of *Natrix* in Sardinia), yielded largely congruent results (Fig. 3; Table 2). This suggests that the reopening of the Strait of Gibraltar played for water snakes an important role as a vicariant event indeed.

Our divergence time estimates are significantly older than those suggested by Guicking et al. (2006, 2008), who used average divergence rates of amino acids and mtDNA sequences assuming linear relationships between divergence times and genetic distances. Guicking et al. (2006, 2008) suggested that differentiation of the genetic lineages of *N. natrix* commenced about 6 Ma. Our mean estimates for the onset of divergence in *N. natrix* are approximately 9.6 Ma (fossil calibration) and 10.6 Ma (palaeogeographic calibration; Table 2: node 1). According to our calculations, Corso-Sardinian grass snakes diverged some 4.3–4.4 Ma ago (both calibrations) from the clade embracing sequences of *N. n. helvetica* and *N. n. lanzai* (Table 2: node 3). Most clades were estimated to be of Upper Miocene age, and only the diversification of peninsular Italian snakes (*N. n. lanzai*) and *N. n. helvetica* (nodes 6 and 7), of Balkan grass snakes (Greece, Hungary, Romania, Slovenia; nodes 11 and 12), and of eastern European and Anatolian/Caucasian grass snakes (Russia, Armenia, Georgia, Turkey; node 14) are likely to be of Pleistocene origin. However, the divergence times of a highly distinct sequence of a *N. n. helvetica* from the southern Alpine part of Switzerland (canton Ticino) and a grass snake from Bulgaria were estimated to predate the Pleistocene (nodes 5 and 13).

Discussion

Due to their pronounced morphological differences, the taxonomic distinctiveness of Corsican and Sardinian grass snakes has never been doubted since their descriptions (Gené 1839; Hecht 1930), and even the parsimonious intraspecific classification system of Thorpe (1979) recognised both as valid subspecies of *Natrix natrix*. Later, using allozyme analyses, Hille (1997) found Sardinian grass snakes to be clearly differentiated, while previous phylogeographic studies did not include specimens from Corsica and Sardinia due to their rareness. Especially Sardinian grass snakes are extremely rare, most probably due to niche competition with the invasive viperine snake *N. maura* (Vanni and Cimmaruta 2010). Here we present for the first time an assessment of the phylogeography of grass snakes including both specimens from Corsica and Sardinia.

According to the geographic distribution of the distinct mitochondrial clades (Fig. 4), the phylogeography of grass snakes matches a common pattern, with each of the three major southern European peninsulas plus Anatolia and Corso-Sardinia harbouring at least one old genetic lineage that often predates the Pleistocene (e.g., Hewitt 2000; Joger et al. 2007; Schmitt 2007; Taberlet et al. 1998). Like in the European pond turtle (*Emys orbicularis*; Fritz et al. 2007; 2009), an Anatolian/Caucasian lineage colonised the vast northeastern part of the range, most probably only in the Holocene, and allied lineages occur on the Balkan Peninsula. A further distinct lineage was recorded from Iran.

As already noted by Guicking et al. (2008), most of the mitochondrial lineages (clades) do not correspond well to currently recognised taxa within the grass snake complex. Based on morphological evidence, Thorpe (1979) proposed that only four subspecies within *N. natrix* should be recognised, viz. *N. n. natrix* (eastern part of the range to the Rhine region), *N. n. helvetica* (western part of the range to the Rhine region, where it hybridises with *N. n. natrix*), *N. n. cetti* (Sardinia), and *N. n. corsa* (Corsica). Many authors, however, continued to distinguish approximately ten additional subspecies that differ mainly in coloration (e.g. Blosat 2008; Gruber 1989; Kabisch 1999; Kreiner 2007), and recently the Corso-Sardinian grass snakes were split off as the

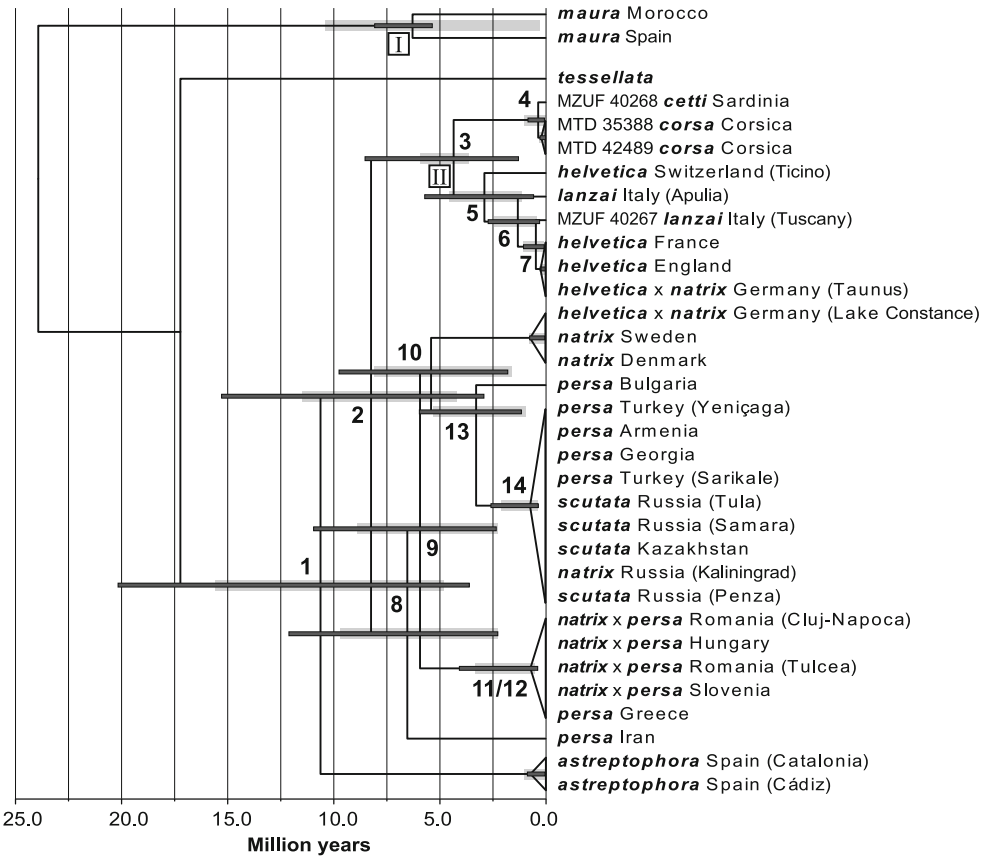

Fig. 3 Estimated split ages of grass snake clades and their 95% HPD intervals (*grey bars*). *Narrow grey bars* are derived from the dating approach using the post-Messinian reopening of the Strait of Gibraltar as age constraint (calibration point I); *wide grey bars*, using the Sardinian fossil node constraint (calibration point II). Numbers along nodes refer to Table 2; see there for exact values. The depicted nodal ages are based on calibration point I

distinct species *N. cetti* (with the subspecies *N. c. cetti* and *N. c. corsa*; Vanni and Cimmaruta 2010).

Most of the traditionally recognised taxa do not match mitochondrial clades, and this is also true for Thorpe's (1979) more inclusive concept for *N. n. natrix* and *N. n. helvetica*. Only the sequences of *N. n. astreptophora* and of the Corso-Sardinian grass snakes constitute two distinct and well-supported clades (Fig. 1) of substantial age (Fig. 3; Table 2). The mean divergence time estimates for *N. n. astreptophora* (9.6 and 10.6 Ma) and *N. cetti* (4.4 and

Table 2 Divergence time estimates (means; in brackets, 95% HPD intervals) for nodes in grass snake phylogeny in million years. Hybrids are subsumed under the subspecies corresponding to their mitochondrial haplotype. Node numbers refer to Fig. 3

Node	Fossil calibration	Palaeogeographic calibration
1 – *astreptophora*+(all other grass snakes)	9.58 (4.80–15.57)	10.61 (3.60–20.15)
2 – Clade A+clade B	7.34 (4.22–11.50)	8.22 (2.91–15.27)
3 – (*cetti*+*corsa*)+(*helvetica*, *lanzai*)	4.44 (3.63–5.93)	4.33 (1.28–8.51)
4 – *cetti*+*corsa*	0.41 (0.04–1.04)	0.35 (0.06–0.83)
5 – *helvetica* [Switzerland]+(other *helvetica*, *lanzai*)	2.88 (1.14–4.56)	2.88 (0.57–5.70)
6 – *lanzai* [Apulia]+(*lanzai* [Tuscany]+remaining *helvetica*)	1.32 (0.42–2.44)	1.31 (0.28–2.71)
7 – *lanzai* [Tuscany]+remaining *helvetica*	0.49 (0.09–1.08)	0.46 (0.06–1.04)
8 – *persa* [Iran]+(all other *persa*, *natrix*, *scutata*)	5.61 (2.25–9.66)	6.52 (2.25–12.10)
9 – *persa* [Greece, Hungary, Romania, Slovenia]+(remaining *persa*, *natrix*, *scutata*)	5.08 (2.26–8.86)	5.92 (2.33–10.93)
10 – *natrix*+(*natrix* [Kaliningrad], *persa* [Armenia, Georgia, Turkey, Bulgaria], *scutata*)	4.53 (1.63–8.09)	5.40 (1.78–9.73)
11 – *persa* [Greece]+*persa* [Hungary, Romania, Slovenia]	1.76 (0.54–3.34)	1.86 (0.37–4.06)
12 – *persa* [Hungary, Romania]+*persa* [Romania, Slovenia]	0.84 (0.20–1.68)	0.88 (0.16–1.93)
13 – *persa* [Bulgaria]+(*natrix* [Kaliningrad], *persa* [Armenia, Georgia, Turkey], *scutata*)	2.86 (0.97–5.31)	3.27 (1.14–5.95)
14 – Basal divergence within clade (*natrix* [Kaliningrad], *persa* [Armenia, Georgia, Turkey], *scutata*)	1.16 (0.39–2.11)	1.22 (0.35–2.58)

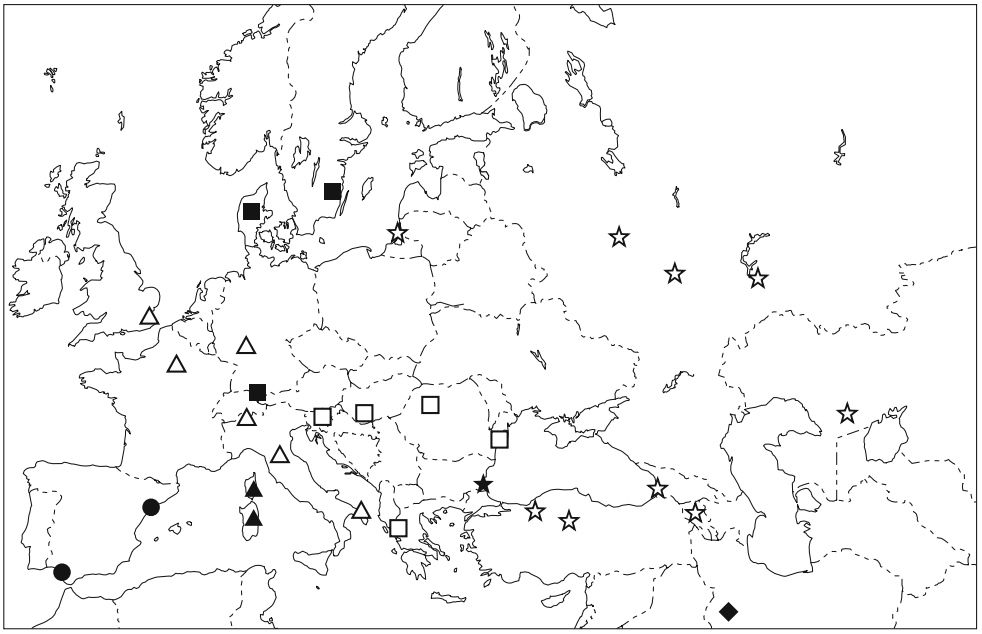

Fig. 4 Geographic distribution of mitochondrial clades in grass snakes. Symbols correspond to Fig. 1

4.3 Ma) suggest a long independent history of the two lineages. The origin of *N. n. astreptophora* could be related with the Alpine orogenesis that has been hypothesized as a major vicariance event for freshwater fishes (Zardoya and Doadrio 1999). Considering the 95% HPD intervals, the divergence estimates for the Corso-Sardinian *N. cetti* agree well with the Zanclean flooding of the Mediterranean Basin (5.33 Ma; García-Castellanos et al. 2009), suggesting that the rising Pliocene sea level triggered their split-off. However, sequences of *N. c. cetti* and *N. c. corsa* are differentiated only weakly (Figs. 1, 2). Their estimated mean divergence times (0.41 and 0.35 Ma) agree well with a hypothetical Pleistocene vicariance event due to the well-known intermittent land connections of Corsica and Sardinia caused by Pleistocene sea level fluctuations (Lambeck et al. 2004).

An unexpected finding of our study was that sequences of *N. natrix* are clearly paraphyletic with respect to *N. cetti* (Fig. 1). The sister group of *N. cetti* is a clade comprising sequences of *N. n. helvetica* and *N. n. lanzai*. There is overwhelming morphological and allozymic evidence for the large-scale hybridisation of *N. n. helvetica* with *N. n. natrix* along the Rhine region (Hille 1997; Thorpe 1979). If these two taxa, being phylogenetically more distantly related than *N. n. helvetica* and *N. cetti* (Fig. 1) and still capable of complete genetic amalgamation, are used as a yardstick, it seems likely that also the allopatric Corso-Sardinian grass snakes possess a similar genetic compatibility. Consequently, the recently proposed species status of Corso-Sardinian grass snakes is no longer tenable under the Biological Species Concept.

Two western German grass snakes cluster either with *N. n. helvetica* or *N. n. natrix*, in agreement with the evident hybridisation of the two subspecies in the Rhine region. However, as already noted by Guicking et al. (2008), sequences of *N. n. helvetica* are paraphyletic with respect to *N. n. lanzai*, suggesting their synonymy. On the other hand, there are deep divergences within this paraphyletic clade (Fig. 1). This could argue for the existence of more than two distinct subspecies, especially when the weak genetic differentiation of the morphologically highly distinct grass snake subspecies from Corsica and Sardinia is considered (Fig. 2). It is obvious that the differences within the *helvetica-lanzai* clade (including a German hybrid *N. n. helvetica* x *natrix* with *helvetica* haplotype) are much more pronounced. The Tuscan haplotype of *N. n. lanzai* is closer to *N. n. helvetica* from England and France and to the German hybrid specimen than to another haplotype of *N. n. lanzai* from Apulia. The haplotype of a *N. n. helvetica* from the Swiss canton Ticino, located south of the Alps, is highly distinct. Together with the considerable age estimates for these lineages, this situation suggests that several distinct microrefuges existed on the Apennine Peninsula during the Pleistocene, as has been shown for quite a number of other species (see the review in Pedall et al. 2011), and that along the southern slope of the Alps was another refuge located that harboured the lineage represented by the haplotype from Ticino. The similarity of the haplotype of our new Tuscan sample to haplotypes from England, France and western Germany implies a Holocene recolonisation of these more northerly regions from a western Italian glacial refuge, as in the case of the green lizard (*Lacerta bilineata*; Böhme et al. 2007) and the Aesculapian snake (*Zamenis longissimus*; Musilová et al. 2010).

Another deeply divergent clade, with estimated mean ages of 4.5 and 5.4 Ma, corresponds to *N. n. natrix* sequences from Denmark and Sweden, and a further hybrid snake

from Germany (Lake Constance). The geographic distribution of this mitochondrial lineage in Central and Northern Europe, intercalated among three other lineages, implies that grass snakes must have survived the last glacial also in a refuge north of the Alps and north of the Balkan Peninsula.

Sequences of the striped subspecies *N. n. persa* occur in four deeply divergent clades (Fig. 1), three of which were also included in the analyses of Guicking (2004) and Guicking et al. (2006, 2008). As Guicking et al. (2008) correctly noted, the placement of sequences of the nominal subspecies *N. n. natrix* (Kaliningrad Oblast, Russia), *N. n. persa* (Armenia, Georgia, Turkey) and *N. n. scutata* (Kazakhstan, Russia) in one clade with shallow divergences suggests that delineation of these subspecies needs revision. The subspecies *N. n. persa* is characterised by longitudinal back stripes, whereas the other two subspecies are unstriped (Kabisch 1999). This allows the conclusion that the striped pattern is merely a coloration morph that occurs in distinct genetic lineages—in other words: "not every striped grass snake is the same".

The distinct endemic lineages in the Balkan Peninsula suggest that at least two glacial refugia were located there.

Conclusions

The delineation of grass snake subspecies needs revision and would profit from a range-wide phylogeography using a denser sampling. Most mitochondrial clades do not match morphologically defined subspecies, suggestive of inappropriate morphological taxon delineation and mitochondrial introgression. The most differentiated mitochondrial clade corresponds to the Iberian subspecies *Natrix natrix astreptophora*, while Corso-Sardinian grass snakes represent the sister group of a paraphyletic assemblage comprising sequences of *N. n. helvetica* and *N. n. lanzai*. Considering that *N. n. helvetica* is known to hybridize extensively with the nominotypical subspecies, the species status of Corso-Sardinian grass snakes is no longer tenable and their status should be reinstated as subspecies of *N. natrix*. The pronounced sequence divergence of the paraphyletic clade embracing *N. n. helvetica* and *N. n. lanzai* suggests the existence of multiple glacial refugia south of the Alps and in the Apennine Peninsula. Western Europe was recolonized from one of these refuges in the Holocene. On each of the other southern European peninsulas and in Anatolia was at least one other glacial refuge located; for the Balkan Peninsula there is evidence for two distinct refugia. In such refugia genetic lineages of often pre-Pleistocene age survived. The estimated divergence times of some grass snake clades are considerable and suggest that the onset of differentiation is related with the Alpine orogenesis. Corso-Sardinian grass snakes are an ancient lineage that seems to have diverged in the Early Pliocene. Another deeply divergent clade of approximately the same age from central and northern Europe suggests the Pleistocene survival of grass snakes north of the Alps.

Acknowledgements Annamaria Nistri allowed sampling of grass snakes from the collection of the Museo di Storia Naturale dell'Università di Firenze. Laboratory work was done by Anke Müller.

References

Arnold, E. N., & Burton, J. A. (1978). *Reptiles and Amphibians of Britain and Europe*. London: Collins.

Benton, M. J., Donoghue, P. C. J., & Asher, R. J. (2009). Calibrating and constraining molecular clocks. In S. B. Hedges & S. Kumar (Eds.), *The Time Tree of Life* (pp. 35–86). Oxford: Oxford University Press.

Blosat, B. (2008). Population status, threats and protection of the grass snake, *Natrix natrix cypriaca* (Hecht, 1930) on Cyprus. *Mertensiella, 17*, 246–271.

Böhme, M. U., Fritz, U., Kotenko, T., Džukić, G., Ljubisavljević, K., Tzankov, N., & Berendonk, T. U. (2007). Phylogeography and cryptic variation within the *Lacerta viridis* complex. *Zoologica Scripta, 36*, 119–131.

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

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland: Sinauer Associates.

de Queiroz, A., Lawson, R., & Lemos-Espinal, J. A. (2002). Phylogenetic relationships of North American garter snakes (*Thamnophis*) based on four mitochondrial genes: how much DNA sequence is enough? *Molecular Phylogenetics and Evolution, 22*, 315–329.

Delfino, M., Bailon, S., & Pitruzella, G. (2011). The Late Pliocene amphibians and reptiles from "Capo Mannu D1 Local Fauna" (Mandriola, Sardinia, Italy). *Geodiversitas, 33*, 357–382.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian Evolutionary Analysis by Sampling Trees. *BMC Evolutionary Biology, 7*, 214.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology, 4*, e88.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution, 39*, 783–791.

Fritz, U., Guicking, D., Kami, H., Arakelyan, M., Auer, M., Ayaz, D., Ayres Fernández, C., Bakiev, A. G., Celani, A., Džukić, G., Fahd, S., Havaš, P., Joger, U., Khabibullin, V. F., Mazanaeva, L. F., Široký, P., Tripepi, S., Valdeón Vélez, A., Velo Antón, G., & Wink, M. (2007). Mitochondrial phylogeography of European pond turtles (*Emys orbicularis*, *Emys trinacris*)—an update. *Amphibia-Reptilia, 28*, 418–426.

Fritz, U., Ayaz, D., Hundsdörfer, A. K., Kotenko, T., Guicking, D., Wink, M., Tok, C. V., Çiçek, K., & Buschbom, J. (2009). Mitochondrial diversity of European pond turtles (*Emys orbicularis*) in Anatolia and the Ponto-Caspian Region: Multiple old refuges, hotspot of extant diversification and critically endangered endemics. *Organisms, Diversity and Evolution, 9*, 100–114.

García-Castellanos, D., Estrada, F., Jiménez-Munt, I., Gorini, C., Fernández, M., Vergés, J., & De Vicente, R. (2009). Catastrophic flood of the Mediterranean after the Messinian salinity crisis. *Nature, 462*, 778–781.

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

Gruber, U. (1989). *Die Schlangen Europas und rund ums Mittelmeer*. Stuttgart: Kosmos.

Guicking, D. (2004). *Molecular phylogeography and evolution of western Palaearctic water snakes (genus* Natrix*, Reptilia)*. Dissertation, Heidelberg University: Combined Faculties for the Natural Sciences and Mathematics.

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.

Gustincich, S., Manfioletti, G., del Sal, G., Schneider, C., & Carninci, C. (1991). A fast method for high-quality genomic DNA extraction from whole human blood. *Bio Techniques, 11*, 298–302.

Hecht, G. (1930). Systematik, Ausbreitungsgeschichte und Ökologie der europäischen Arten der Gattung *Tropidonotus* (Kuhl) H. Boie. *Mitteilungen aus dem Zoologischen Museum Berlin, 16*, 244–393.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature, 405*, 907–913.

Hille, A. (1997). Biochemical variation between populations of the western and eastern grass snake (*Natrix natrix*) from the transition zone in Nordrhein-Westfalen, Germany. In W. Böhme, W. Bischoff, & T. Ziegler (Eds.), *Herpetologia Bonnensis* (pp. 177–184). Bonn: Societas Herpetologica Europaea.

Joger, U., Fritz, U., Guicking, D., Kalyabina-Hauf, S., Nagy, Z. T., & Wink, M. (2007). Phylogeography of western Palaearctic reptiles—Spatial and temporal speciation patterns. *Zoologischer Anzeiger, 246*, 293–313.

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

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

Kreiner, G. (2007). *Schlangen Europas*. Frankfurt/M.: Chimaira.

Lacepède, B.-G.-É. (1789). *Histoire naturelle des quadrupèdes ovipares et des serpens, vol. 2 (Histoire naturelle des serpens)*. Paris: Hôtel de Thou.

Lambeck, K., Antonioli, F., Purcell, A., & Silenzi, S. (2004). Sea-level change along the Italian coast for the past 10,000 yr. *Quaternary Science Reviews, 23*, 1567–1598.

Lanza, B. (1988). Hypothèses sur les origines de la faune herpétologique Corse. *Bulletin d'Ecologie, 19*, 163–170.

Laurenti, J. N. (1768). *Specimen Medicum, Exhibens Synopsin Reptilium Emendatam cum Experimentis circa Venena et Antidota Reptilium Austriacorum*. Vienna: Trattnern.

Lawson, R., Slowinski, J. B., Crother, B. I., & Burbrink, F. T. (2005). Phylogeny of the Colubroidea (Serpentes): new evidence from mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution, 37*, 581–601.

Linnaeus, C. (1758). *Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis, Synonymis, Locis. Tomus I. Regnum Animale. Editio Decima*. Stockholm: Laurentius Salvius.

Linnaeus, C. (1766). *Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis, Synonymis, Locis. Tomus I. Regnum Animale. Editio Duodecima*. Stockholm: Laurentius Salvius.

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

Musilová, R., Zavadil, V., Marková, S., & Kotlík, P. (2010). Relics of the Europe's warm past: Phylogeography of the Aesculapian snake. *Molecular Phylogenetics and Evolution, 57*, 1245–1252.

Nylander, J. A. A. (2004). *MrMODELTEST, v2.* Uppsala: Evolutionary Biology Centre, Uppsala University, http://www.abc.se/~nylander/ [accessed 29 January 2011].

Pallas, P. S. (1771). *Reise durch verschiedene Provinzen des Rußischen Reichs. Erster Theil*. St. Petersburg: Kayserliche Academie der Wissenschaften.

Pallas, P. S. (1814). *Zoographia Rosso-Asiatica, sistens Omnium Animalium in Extenso Imperio Rossico et Adjacentibus Maribus Observatorum Recensionem, Domicilia, Mores et Descriptiones, Anatomen atque Icones Plurimorum. Volumen Tertium*. Petropolis: Caesarea Academia Scientiarum.

Pedall, I., Fritz, U., Stuckas, H., Valdéon, A., & Wink, M. (2011). Gene flow across secondary contact zones of the *Emys orbicularis* complex in the Western Mediterranean and evidence for extinction and re-introduction of pond turtles on Corsica and Sardinia (Testudines: Emydidae). *Journal of Zoological Systematics and Evolutionary Research, 49*, 44–57.

Rambaut, A., & Drummond, A. J. (2007). *TRACER v1.4.* Auckland: Bioinformatics Institute, University of Auckland, http://beast.bio.ed.ac.uk/Tracer [accessed 1 February 2011].

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 12*, 1572–1574.

Sabaj Pérez, M. H. (2010). *Standard symbolic codes for institutional resource collections in herpetology and ichthyology: an Online Reference. Version 1.5 (4 Oct 2010).* Washington, DC: American Society of Ichthyologists and Herpetologists, http://www.asih.org/ [accessed 15 July 2011].

Schmitt, T. (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology, 4*, 11.

Seoane, V. L. (1884). *Identidad de* Lacerta schreiberi *(Bedriaga) y* Lacerta viridis *var.* gadovi *(Boulenger) e investigaciones herpetológicas de Galicia*. Coruña: Vicente Abad.

Silvestro, D., & Michalak, I. (2011). raxmlGUI: a graphical front-end for RAxML. *Organisms, Diversity and Evolution*. doi:10.1007/s13127-011-0056-0.

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

Swofford, D. L. (2002). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), ver. 4*. Sunderland: Sinauer Associates.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G., & Cosson, J.-F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology, 7*, 453–464.

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.

Vanni, S., & Cimmaruta, R. (2010). *Natrix cetti* Gené, 1839. In C. Corti, M. Capula, L. Luiselli, E. Razzetti, & R. Sindaco (Eds.), *Fauna d'Italia, Reptilia* (pp. 541–547). Bologna: Calderoni.

Zardoya, R., & Doadrio, I. (1999). Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *Journal of Molecular Evolution, 49*, 227–237.