Molecular Phylogenetics and Evolution 180 (2023) 107674

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogenomics of *Anguis* and *Pseudopus* (Squamata, Anguidae) indicates Balkan-Apennine mitochondrial capture associated with the Messinian event

Václav Gvoždík [a,b,*], Tadeáš Nečas [a,c], Daniel Jablonski [d], Emily Moriarty Lemmon [e], Alan R. Lemmon [f], David Jandzik [d], Jiří Moravec [b]

[a] Institute of Vertebrate Biology of the Czech Academy of Sciences, Brno, Czech Republic
[b] National Museum, Department of Zoology, Prague, Czech Republic
[c] Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic
[d] Department of Zoology, Faculty of Natural Sciences, Comenius University in Bratislava, Ilkovičova 6, Mlynská dolina, 842 15 Bratislava, Slovakia
[e] Florida State University, Department of Biological Sciences, Tallahassee, USA
[f] Florida State University, Department of Scientific Computing, Dirac Science Library, Tallahassee, USA

ARTICLE INFO

Keywords:
Evolutionary history
Glass lizard
Introgression
Messinian salinity crisis
Mitonuclear discordance
Slow worm

ABSTRACT

A dated phylogenetic hypothesis on the evolutionary history of the extant taxa of the Western Palearctic lizards *Anguis* and *Pseudopus* is revised using genome-wide nuclear DNA and mitogenomes. We found overall concordance between nuclear and mitochondrial DNA phylogenies, with one significant exception – the Apennine *A. veronensis*. In mitochondrial DNA, this species forms a common clade with the earliest diverging lineage, the southern Balkan endemic *A. cephallonica*, while it clusters together with *A. fragilis* in nuclear DNA. The nuclear phylogeny conforms to the morphology, which is relatively similar between *A. veronensis* and *A. fragilis*. The most plausible explanation for the mitonuclear discordance is ancient mitochondrial capture from the Balkan ancestor of *A. cephallonica* to the Apennine population of the *A. fragilis-veronensis* ancestor. We hypothesize that this capture occurred only in a geographically restricted population. The dating of this presumed mitochondrial introgression and capture coincides with the Messinian event, when the Balkan and Apennine Peninsulas were presumably largely connected. The dated nuclear phylogenomic reconstruction estimated the divergence of *A. cephallonica* around 12 Mya, while the sister clade representing the *A. fragilis* species complex consisting of the sister species *A. fragilis–A. veronensis* and *A. colchica–A. graeca* further diversified around 7 Mya. The depth of nuclear divergence among the evolutionary lineages of *Pseudopus* (0.5–1.2 Mya) supports their subspecies status.

1. Introduction

The Western Palearctic lacertilians slow worms (*Anguis*) and glass lizards (*Pseudopus*; also known as sheltopusiks, previously in the genus *Ophisaurus*) of the family Anguidae, subfamily Anguinae, with the snake-like legless body form were traditionally considered to represent a single species in their respective genera, *A. fragilis* and *P. apodus* (Dely, 1981; Obst, 1981). Another species of *Anguis*, *A. cephallonica* (Peloponnese slow worm), from the Peloponnese Peninsula and the nearby Ionian Islands in the southern Balkans was taxonomically distinguished in the 1990s (Grillitsch and Cabela, 1990). In the last decade, *A. fragilis* was shown to contain additional cryptic diversity; the taxon was first divided into three species, the widespread western *A. fragilis*, the widespread eastern *A. colchica*, and the southwestern Balkan endemic *A. graeca* (Gvoždík et al., 2010). Later, another, apparently ancient species from the Apennine (Italian) Peninsula and southeastern Mediterranean France, *A. veronensis*, was recognized (Gvoždík et al., 2013). The phylogenetic relationships among the species of the genus *Anguis* were not fully resolved (Gvoždík et al., 2013), but the latter species, the Italian slow worm, was found to have a deep divergence in mitochondrial DNA (mtDNA) with presumably a sister relationship to the Peloponnese slow worm (Thanou et al., 2014, 2021; Jablonski et al., 2016). However, the signal from one of the two nuclear genes studied also suggested a possible close relationship to *A. fragilis*, as did the morphology (Gvoždík et al., 2013). Based on morphology, the Italian populations of slow worms were believed to represent the

* Corresponding author at: Institute of Vertebrate Biology of the Czech Academy of Sciences, Research Facility Studenec, Studenec 122, 675 02, Czech Republic.
E-mail address: vaclav.gvozdik@gmail.com (V. Gvoždík).

https://doi.org/10.1016/j.ympev.2022.107674
Received 12 January 2022; Received in revised form 19 September 2022; Accepted 7 December 2022
Available online 18 December 2022

nominotypical *A. fragilis fragilis* (Völkl and Alfermann, 2007) in the past, i.e. what later became *A. fragilis* sensu stricto. *Anguis colchica* is the only species of the genus with recognized subspecies: *A. c. colchica* distributed in the Caucasus region, *A. c. orientalis* in the southern Caspian region, and *A. c. incerta* in most of the European range of the species, spatially separated from the Caucasus–Caspian populations (Gvoždík et al., 2010, 2013; Jablonski et al., 2021b). In addition, another unnamed evolutionary lineage of *A. colchica*, the so-called Pontic clade of *A. colchica*, was identified in the southeastern Balkans and northwestern Anatolia (Jablonski et al., 2016, 2021b). In contrast, mtDNA variation of the second geographically widespread species, *A. fragilis*, was found to be surprisingly low (Gvoždík et al., 2010, 2013; Jablonski et al., 2016; 2017). Diversification of the extant slow worm species was estimated to originate in the Late Miocene, whereas of the subspecies-level lineages during the Pliocene-Pleistocene transition (Gvoždík et al., 2010). However, not much is known about the fossil diversity of *Anguis*, with majority of the fossils, mostly from the Pleistocene, attributed to *A.* cf. *fragilis* or the *A. fragilis* species complex (Klembara and Rummel, 2018; Villa and Delfino, 2019). The earliest known fossil representative of the genus *Anguis*, the approximately 20 My old *A. rarus*, was described from Germany (Klembara and Rummel, 2018). The current distributions of individual species, which are exclusively or predominantly European, were recently revised (Jablonski et al., 2021b).

The glass lizard of the genus *Pseudopus* is much larger in body size than *Anguis* (up to 135 cm vs 60 cm total length), distributed from southeastern Europe to Central Asia. *Pseudopus* and *Anguis* are usually considered to have sister relationship (Macey et al., 1999; Pyron et al., 2013; Lavin and Girman, 2019) with origin between 18.4 and 24.5 Mya (Lavin and Girman, 2019). However, molecular species-tree analysis of the subfamily Anguinae also showed *Pseudopus* as a sister lineage to the Asiatic genus *Dopasia*, along with the American genus *Ophisaurus*, as an alternative phylogenetic scenario (Lavin and Girman, 2019), which is also partially consistent with morphology-based phylogenetic reconstruction (Klembara et al., 2014). Another morphology-based alternative shows *Anguis* as a sister lineage to *Ophisaurus* (including *Dopasia*; the two genera are sometimes synonymized) and *Pseudopus* is placed at a more distant position (Klembara et al., 2019). This means that the intergeneric relationships within the Anguinae subfamily are not yet reliably known, although a scenario consisting of two clades (Western Palearctic with *Anguis* and *Pseudopus* as sister genera and East Asiatic-North American) seems more likely (Lavin and Girman, 2019). The earliest fossils of *Pseudopus* (*P. ahnikoviensis*) are known from the Early Miocene about 20 Mya (Klembara and Rummel, 2018; Villa and Delfino, 2019), with older records of the Anguinae from the Oligocene likely attributable to other, presumably extinct genera, such as *Ophisauromimus* (Čerňanský et al., 2016). Although the genus *Pseudopus* has a relatively rich fossil record and diversity of extinct species since the Early Miocene (Klembara and Rummel, 2018; Villa and Delfino, 2019), only three evolutionary lineages at the subspecies level are currently recognized within a single species, *P. apodus apodus* from the Caucasus to Central Asia, *P. a. thracius* from the Balkans and western Anatolia, and the recently described *P. a. levantinus* from the Levant (Jandzik et al., 2018; Jablonski et al., 2021a). These three subspecies probably diversified during or shortly before the Pleistocene (Jandzik et al., 2018), although an alternative scenario of diversification in the Late Miocene to Early Pliocene was also presented (Lavin and Girman, 2019). This latter scenario would be more consistent with diversification at the species rather than subspecies level, therefore a new time-calibrated phylogeny at the genomic scale could help resolve this inconsistency.

The aim of this study is to (1) build a time-calibrated species tree of extant taxa in the genera *Anguis* and *Pseudopus* using a robust genome-wide molecular dataset, (2) compare the nuclear DNA phylogenomic reconstruction with the mitogenomic reconstruction and investigate potential discordances, and (3) provide a unifying hypothesis for their evolutionary history.

2. Material and methods

2.1. Sampling design

The genus *Anguis* was represented by 10 specimens of the all five recognized species. The samples were selected to cover all subspecies of *A. colchica* and one as yet scientifically unnamed distinct evolutionary lineage of this species. Two widely distributed taxa, *A. fragilis* and *A. colchica incerta*, were represented by two geographically distant samples. One of them originated from the vicinity of the secondary contact zone between *A. fragilis* and *A. colchica* in the Czech Republic (Jablonski et al., 2021b) in order to test in these morphologically difficult-to-distinguish species (Benkovský et al., 2021) their putative phylogenetic divergence within a geographically relatively small area. The genus *Pseudopus* was represented by three samples, each belonging to a different subspecies of *P. apodus*. One outgroup was represented by the Oriental glass-lizard genus *Dopasia* (*D. sokolovi*). For the geographic origin of the samples, see Table 1 and Fig. 1A. Most genetic (tissue) samples were not documented by museum voucher specimens, with the exception of *A. colchica orientalis* (NMP-P6V 72678), *P. apodus levantinus* (TAU-R 16928) and *Dopasia sokolovi* (NCSM 77336); museum abbreviations follow Sabaj (2022).

2.2. Sequence capture and assembly of mitogenomes

DNA was extracted from tissue samples using various commercial spin column-based kits according to the manufacturers' protocols. Genomic sequence capture data were generated for 14 individuals, including one outgroup, using the anchored hybrid enrichment methodology described by Lemmon et al. (2012). See Supplementary material for technical details. The probes used for the anchored hybrid enrichment successfully captured 387 loci, but four loci contained a high amount of missing data, resulting in 383 loci being analyzed. Altogether, 339 of the alignments contained all 14 individuals, while 44 alignments had one (15 alignments) or more individuals missing (14 alignments with two to three individuals missing; 14 alignments with four to six individuals missing; one alignment with seven individuals missing). The full concatenated dataset comprised 687,027 aligned bp and contained 4.6 % of missing characters.

A total of 14 mitogenomes were assembled in Geneious Prime v2021.2.2 (https://www.geneious.com) from raw anchored hybrid enrichment Illumina reads using available annotated mitogenomes from

Table 1
List of samples and their geographic origin.

Taxon / Lineage	Country	Locality	Lat (°N)	Lon (°E)
Anguis cephallonica	Greece	Gialova	36.95	21.70
Anguis colchica colchica	Georgia	Telavi	41.92	45.49
Anguis colchica incerta CZ	Czech Rep.	Ostravice	49.52	18.38
Anguis colchica incerta LT	Lithuania	Marcinkonys	54.04	24.44
Anguis colchica orientalis	Iran	Motalla Sara-ye Lemir	38.20	48.87
Anguis colchica "Pontic"	Bulgaria	Sinemorec	42.06	27.97
Anguis fragilis CZ	Czech Rep.	Želízy	50.42	14.47
Anguis fragilis ES	Spain	Torla-Ordesa	42.62	−0.11
Anguis graeca	Greece	Mornos River	38.49	22.06
Anguis veronensis	Italy	Roccagnano	43.38	12.11
Pseudopus apodus apodus	Georgia	Dedop'lis Tskaro	41.43	46.10
Pseudopus apodus levantinus	Israel	Hare Gilboa'	32.45	35.43
Pseudopus apodus thracius	Albania	Zogaj	42.07	19.39
Dopasia sokolovi	Vietnam	Bidoup-Nui Ba	12.18	108.68

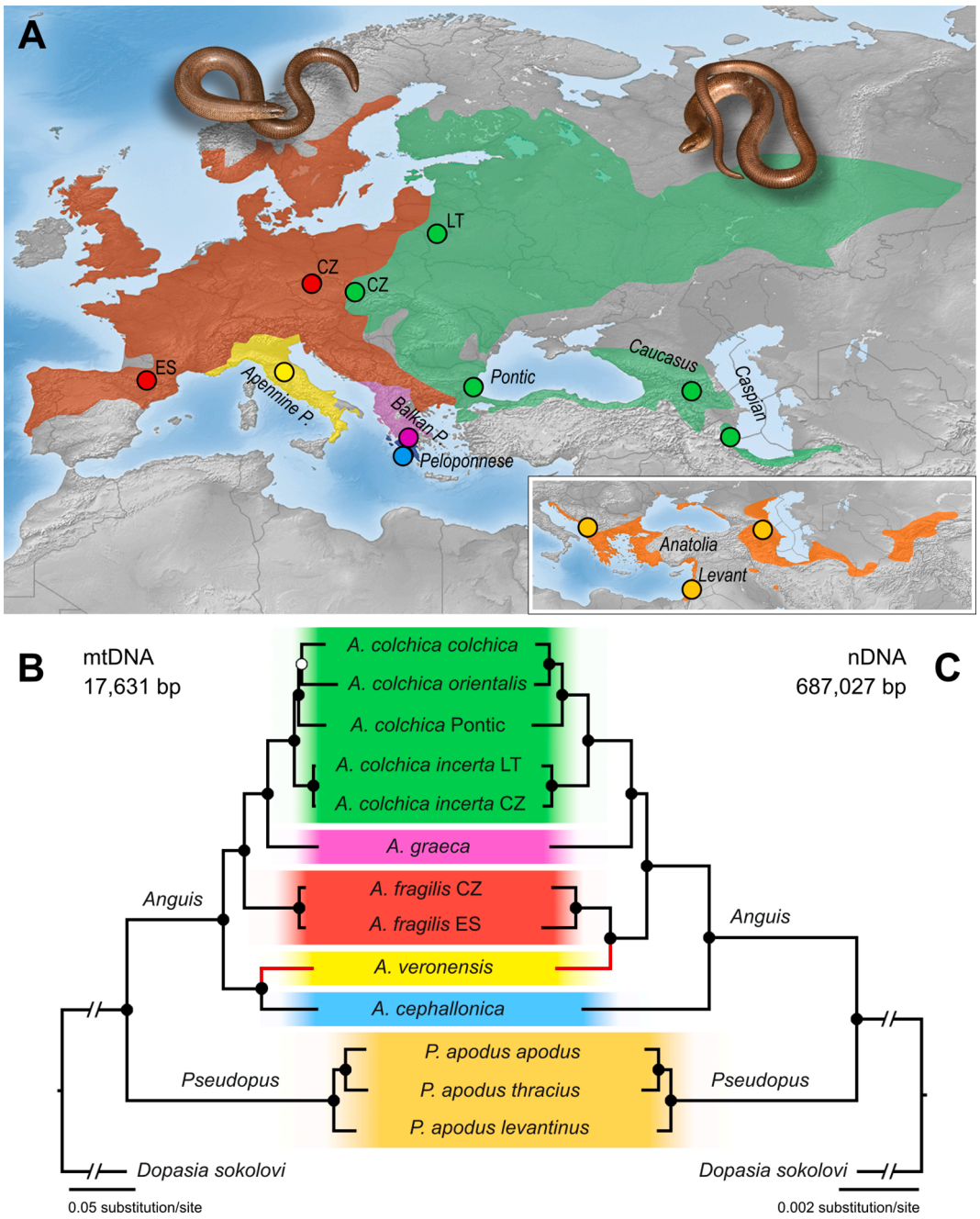

Fig. 1. Distribution ranges and sampling of *Anguis* and *Pseudopus*, and phylogenomic reconstructions of their evolutionary history. (A) Range map and sampling sites of *Anguis*, and *Pseudopus* in the inset map. (B) Mitochondrial DNA phylogenomic tree (BI); and (C) nuclear DNA phylogenomic tree (BI). Note the mitonuclear discordance of the Italian slow worm *A. veronensis* (red branches). All nodes received full support in all analyses (black dots), except for the *A. colchica colchica-A. c. orientalis* clade in the mitogenomic tree (white dot). Photos (not to scale): *A. fragilis* (left), *A. colchica* (right). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the GenBank online database as mapping references (all five *Anguis* species and *Dopasia gracilis* were used as references; *Pseudopus* was assembled against both *Anguis* and *Dopasia*). The final assembly resulted in a 17,631 bp-long alignment, aligned using MAFFT v7 (Katoh and Standley, 2013).

2.3. Standard phylogenetic analyses

2.3.1. nDNA and mtDNA

MrBayes v3.2.7 (Ronquist et al., 2012) and IQ-TREE v2 (Minh et al., 2020) were used to perform Bayesian inference (BI) and maximum likelihood (ML) partitioned analyses, using the best-fit models of nucleotide evolution selected by the model selection module of IQ-TREE v2, ModelFinder (Kalyaanamoorthy et al., 2017) using the Bayesian information criterion. Nuclear genome (nDNA) analyses were performed on a concatenated dataset consisting of a total of 383 partitions corresponding to the 383 loci. Mitogenome analyses were performed on a concatenated dataset consisting of a total of 64 partitions (separate models were used for each codon position in protein-coding genes): 12S and 16S ribosomal RNA (*12S* and *16S*), NADH dehydrogenase subunits 1–6 (*ND1–6*), including 4L (*ND4L*), cytochrome *c* oxidase subunits 1–3 (*COX1–3*), cytochrome *b* (*CYTB*), ATP synthase subunits 6 and 8 (*ATP6* and *8*), control region (*CR*), and 22 transfer RNAs (*tRNAs*). For the distribution of partitions and selected substitution models, see

Supplementary material, Tables S1–S4. BI analyses of both nuclear and mitogenomic datasets were run in duplicates for 10 million generations each, sampling every 5000th generation, with the first 25 % as a burn-in after checking for convergence of parameter estimates and values of the effective sample size (ESS) in Tracer v1.7 (Rambaut et al., 2018). ML analyses of both datasets were run with 1000 ultrafast bootstrap pseudoreplicates (UFBoot2; Hoang et al., 2018).

2.3.2. Testing possible effect of long branch attraction in mtDNA

In order to identify whether the long branch attraction (LBA; Felsenstein, 1978) artifact could affect our mtDNA phylogenetic reconstruction—relatively long branches of *A. cephallonica* and *A. veronensis* may have been artificially attracted together (see Results and discussion)—we performed the following investigations. First, we added the available mitogenomes of *A. cephallonica* and *A. veronensis* from GenBank to confirm the positions of our sequences. We then analyzed the data using the ML method (IQ-TREE v2, UFBoot2: 1000 pseudoreplicates) after (i) adding an additional outgroup from another family within the same clade Neoanguimorpha (Helodermatidae, *Heloderma suspectum*) and alternatively also the sister subfamily Gerrhonotinae (*Abronia graminea*); (ii) removal of third codon positions of protein-coding loci and ambiguously aligned positions of non-protein-coding loci using Gblocks v0.91b (Castresana, 2000); and (iii) division into 17 individual mitochondrial loci (*tRNA*s treated together as 'one locus'), including the *Heloderma* outgroup in the latter two investigations. These examinations, and especially the second (elimination of rapidly evolving and potentially saturated positions) could help to reveal the potential effect of the LBA artifact (Bergsten, 2005).

2.4. Divergence dating

Chronograms were constructed within the coalescent-based framework using StarBeast3 (Douglas et al., 2022) and Standard BEAST templates of BEAST 2 (Bouckaert et al., 2019) for the multilocus nDNA and genome-scale mtDNA data, respectively (mtDNA usually behaves as a single non-recombining locus). Because we were unable to perform the StarBeast3 analysis using all 383 loci, we reduced the dataset by omitting loci containing <11 terminals (out of 14), leaving 368 loci for the StarBeast3 analysis. Substitution models were set up for each locus/partition as for the BI analyses (Supplementary material, Tables S1, S3). In BEAST 2, several preliminary control analyses were conducted with different settings and strict clocks were rejected using the coefficient of variation for the clock rate, therefore the optimized relaxed clock (Zhang and Drummond, 2020; Douglas et al., 2021) was used in both StarBeast3 and Standard BEAST analyses. The birth–death model as the tree prior was used in both analyses. Time calibration priors were set for the MRCA of *Anguis* and *Pseudopus* because the phylogenetic position and divergence dating of *Dopasia* is ambiguous (cf. Klembara et al., 2014, 2019; Lavin and Girman, 2019). The nuclear DNA phylogenetic reconstruction estimated the existence of MRCA of *Anguis* and *Pseudopus* between 18.4 and 24.5 Mya, with a mean across different analyses of around 19–20 Mya (Lavin and Girman, 2019). The earliest known fossils that can be unambiguously assigned to either genus are about 20 My old (Klembara and Rummel, 2018; Villa and Delfino, 2019; J. Klembara, 2021, in litt.), suggesting that the origins of both genera are likely slightly older than 20 Mya. The calibrations of Lavin and Girman (2019) tended to be a bit conservative, working with perhaps slightly younger ages. Comparing the age estimate of the MRCA of the Anguinae and Gerrhonotinae anguid lineages by Lavin and Girman (2019) with the estimate in the phylogenomic study of squamates by Burbrink et al. (2020), the latter study estimated an age slightly older, by about 3 My. Therefore, and considering the earliest *Anguis* and *Pseudopus* fossils (*A. rarus*, *P. ahnikoviensis*) with an approximate age of 20 Mya, we set the calibration for the MRCA of *Anguis-Pseudopus* to the interval 21–25 Mya (mean 23.0 ± 1.0 Mya; normal distribution). The nDNA analysis also included *Dopasia*, and the *Anguis-Pseudopus* sister relationship was not constrained given the uncertainty as to whether these two genera represent sister taxa. The mtDNA analysis was performed without *Dopasia* because preliminary runs with *Dopasia* showed that this outgroup taxon is problematic, placed in a deeply divergent position in the mtDNA. The dating analyses were run in duplicates with different starting seeds, each for 100 million generations and saving each 50,000th generation in the nuclear dataset (StarBeast3), and for 50 million generations and saving each 25,000th generation in the mitogenomic dataset (Standard BEAST). Appropriate burn-in values were identified and ESS values were checked using Tracer v1.7 (Rambaut et al., 2018). Posterior probability distributions of trees were visualized as "cloudograms" in DensiTree v2.2 (Bouckaert and Heled, 2014). Post-burn-in samples were combined in the LogCombiner module of BEAST 2, and the TreeAnnotator module was used to infer the final species tree as a maximum clade credibility tree (MCCT) with node ages calculated as heights of common ancestors and confidence intervals with 95 % highest posterior density (HPD).

3. Results and discussion

3.1. Mitonuclear discordance

The BI and ML produced the same tree topologies based on the concatenated partitioned analyses (Fig. 1B–C). All nodes received full support in all analyses (BI posterior probability/ML bootstrap: 1.00/100), except for the Caucasus-Caspian clade *Anguis colchica colchica-A. c. orientalis* in the mitogenomic tree (0.75/51). The inferred mitogenomic phylogeny is consistent with earlier studies based on shorter mtDNA fragments (e.g., Gvoždík et al., 2013; Jablonski et al., 2016). The topology of the nuclear-genomic phylogeny is concordant with the topology of the mitogenomic phylogeny, with one important exception; i.e. the position of *A. veronensis*, which is a sister lineage to *A. fragilis* in nDNA, as opposed to being a sister lineage to *A. cephallonica* in mtDNA. However, this previously unexpected phylogenetic position of *A. veronensis* (close relationship with *A. fragilis* in nDNA) is in good agreement with the traditional morphology-based systematic approach, in which Italian slow worms were considered to belong to the nominotypical subspecies *A. fragilis fragilis*, i.e., the present-day *A. fragilis* sensu stricto (Völkl and Alfermann, 2007).

The investigations of whether the sister (but divergent) relationship of *A. cephallonica* and *A. veronensis* in mtDNA might be an artifact did not provide support for LBA being responsible for the inferred mtDNA topology. None of the three LBA-focused investigations (see Material and methods) inferred a topology supporting the sister relationship between *A. veronensis* and *A. fragilis*, as would be expected based on nDNA. Instead, all investigations based on mitogenome-scale data returned essentially the same topology (Supplementary material, Figs. S1–S3) as the original BI and ML analyses (Fig. 1B). ML analyses of the 17 individual mitochondrial loci inferred the *A. cephallonica–A. veronensis* clade 9 times with high support; 4 times the clade was inferred in the maximum likelihood tree, albeit without high support; 4 times the clade was not inferred but the two species were still phylogenetically close; whereas the *A. fragilis–A. veronensis* clade was never inferred at any mtDNA locus (Supplementary material, Fig. S4). Moreover, simulations have shown that ML phylogenetic analyses of nucleotide alignments above 10,000 bp (our case) should infer a true tree, while the frequency of LBA trees should be close to zero (Kapli et al., 2020).

However, it must be acknowledged that our sample size is limited and testing the possible effect of the LBA artifact in the future may yield a different conclusion. Allowing that LBA may have potentially affected the topology of the mtDNA phylogenetic tree, we still need to take into account the deep divergence in mtDNA between *A. fragilis* and *A. veronensis*. We hypothesize that if LBA affected the mtDNA topology, the alternative topology would probably show a polytomy, i.e. an unresolved relationship for *A. cephallonica*, *A. fragilis* and *A. veronensis* (own unpublished analyses with a tree constrained for the monophyly of

A. fragilis-A. veronensis). Under these circumstances, alternative hypotheses to the one presented below to explain the divergent position of *A. veronensis* in mtDNA would be either a highly accelerated substitution rate of mtDNA in this species relative to its sister species *A. fragilis* (which would be a very rare case within vertebrate sister species), or mitochondrial capture from a ghost lineage (i.e., an extinct lineage) whose biogeographic origin would be unknown. If such a ghost lineage originated from the Apennine Peninsula, our historical biogeographical hypothesis (Balkan-Apennine mitochondrial capture, see below) would not be valid. However, pre-Pleistocene slow-worm fossils from the Apennine Peninsula are not known (Klembara and Rummel, 2018; Villa and Delfino, 2019). Alternatively, the ghost lineage could originate from the Balkan Peninsula, which seems to be more plausible as the Balkans is the center of the slow-worm diversity. In such case, our biogeographic hypothesis would be valid, except that it would include a ghost lineage instead of the *A. cephallonica* ancestor. We consider these alternative hypotheses to be possible, although in the light of the present data less likely than our main hypothesis outlined below (section 3.2.3).

3.2. Evolutionary history

3.2.1. Comparison of the results of phylogenetic reconstructions

The topology of the chronograms (Fig. 2) is consistent with that of the undated phylogenetic trees (Fig. 1B–C) and received full support at all but one node (*A. colchica colchica-A. c. orientalis* in mtDNA; posterior probability 0.92). The divergence dating was generally consistent when comparing nDNA and mtDNA analyses (Table 2), although HPD intervals are rather wide under the relaxed clock model. The unconstrained nDNA analysis recovered the monophyly of *Anguis* and *Pseudopus*, which further supported the two genera as sister lineages (Macey et al., 1999; Pyron et al., 2013; Lavin and Girman, 2019). However, our sampling was not designed to address this question, so this result should still be taken with caution, as alternative phylogenetic relationships have been inferred in earlier studies (Klembara et al., 2014, 2019; Lavin and Girman, 2019, in part). The mitogenomic data tended to estimate slightly older ages of splits in some terminal branches

Table 2
Times to most recent common ancestors (MRCAs) of the extant taxa and lineages of *Anguis*, *Pseudopus* and the outgroup *Dopasia sokolovi*. Means in bold, 95% HPD interval in parentheses. The asterisk indicates the calibration point set at 23 (21–25) Mya.

MRCA	Nuclear genomic	Mitogenomic
Anguis cephallonica / *A. fragilis* species complex[+]	**11.98** (7.24–17.13)	**8.99** (6.49–11.64)[+]
Anguis cephallonica / *A. veronensis* (mtDNA)	N/A	**5.02** (3.20–6.91)
Anguis fragilis species complex	**6.73** (3.65–10.12)	N/A
Anguis fragilis / *A. colchica-graeca* (mtDNA)	N/A	**6.92** (4.98–9.06)
Anguis graeca / *A. colchica*	**4.43** (2.06–7.25)	**4.81** (3.34–6.31)
Anguis fragilis / *A. veronensis*	**2.74** (0.82–5.11)	N/A
Anguis fragilis CZ / *A. fragilis* ES	**0.39** (<0.01–0.97)	**0.58** (0.34–0.83)
A. colchica	**2.98** (1.35–4.98)	**2.65** (1.89–3.57)
A. colchica South-East	**2.37** (1.06–4.08)	**2.28** (1.54–3.05)
Anguis colchica colchica / *A. colchica orientalis*	**1.39** (0.37–2.56)	**2.04** (1.33–2.73)
Anguis colchica incerta LT / *A. c. incerta* CZ	**0.16** (<0.01–0.46)	**0.24** (0.12–0.35)
Pseudopus apodus	**1.15** (0.32–2.19)	**2.61** (1.77–3.68)
Pseudopus apodus apodus / *P. apodus thracius*	**0.52** (0.11–1.15)	**1.67** (1.08–2.41)
Anguis / *Pseudopus* *	**22.92** (21.05–24.98)	**22.91** (21.15–25.02)
Anguis-Pseudopus / *Dopasia*	**27.61** (22.35–35.54)	–

[+] In the mitogenomic analysis for topology ((*cephallonica*, *veronensis*), (*fragilis*, (*colchica*, *graeca*))).

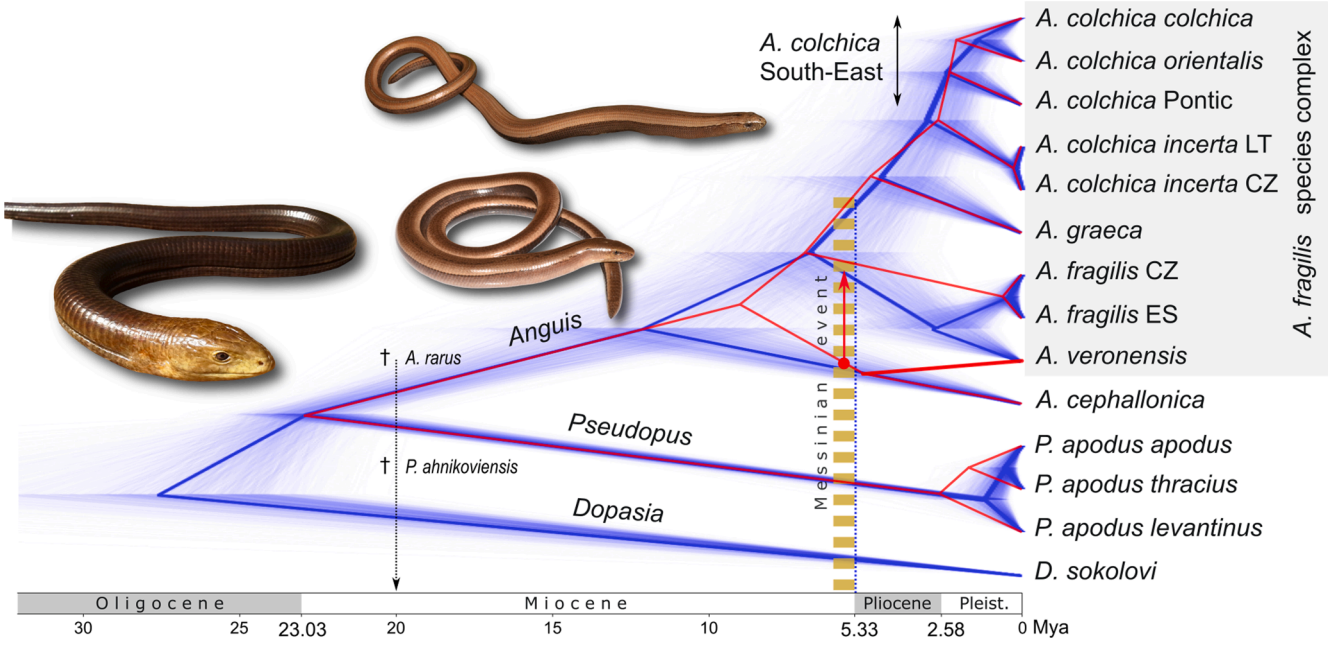

Fig. 2. Chronogram of extant taxa and lineages of *Anguis* and *Pseudopus*. The cloudogram of the post-burn-in samples of the nuclear genomic dataset (blue trees, root canal in bold; 687,027 bp) and the maximum clade credibility tree (MCCT) of the mitochondrial DNA phylogeny on the top (thin red tree; 17,631 bp). The earliest fossil records of *Anguis* (*A. rarus*) and *Pseudopus* (*P. ahnikoviensis*), both of similar age approximately 20 Mya (black dotted arrow), were considered for the calibration of the molecular clock set at 23 (21–25) Mya. All nodes of the respective MCCTs received full support, except for the sister relationship of *A. colchica colchica* and *A. c. orientalis* in the mitogenomic tree. The nuclear and mitochondrial topologies are congruent except for the position of *A. veronensis* (bold red branch in the mtDNA tree). The red arrow indicates the Messinian origin of the assumed mitochondrial capture from the ancestor of *A. cephallonica* to a geographically restricted marginal population of the ancestor of *A. fragilis-veronensis*. The ochre dashed bar indicates the Messinian salinity crisis, the blue dotted line indicates the Zanclean flood. Photos (from top to bottom; not to scale): *A. veronensis*, *A. cephallonica*, *P. apodus*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(e.g. *Pseudopus*), which is probably due to the short coalescence time of mtDNA compared to the incomplete lineage sorting that is commonly present in nDNA in evolutionarily young diversifications (Zheng et al., 2011). Interestingly, the relatively deep divergence between the two *A. fragilis* samples (Czech Republic vs Spain) inferred from the nDNA concatenated analyses (Fig. 1C) was not reflected in the coalescent-based analysis (Fig. 2). Accumulated substitutions, which played a significant role in the concatenated analysis, did not significantly affect the analysis of individual loci in the coalescent framework in this case. We assume that this was an effect of incomplete lineage sorting. The second taxon represented by two samples, *A. colchica incerta* (Lithuania vs Czech Republic), was similarly shallow in divergence in both the concatenated and coalescent-based analyses.

3.2.2. *Anguis cephallonica and A. fragilis species complex*

The earliest division within the genus *Anguis* corresponds to the origin of *A. cephallonica*, a Peloponnese endemic, dating to about 12 Mya at the beginning of the Late Miocene. The Peloponnese slow worm is the most morphologically distinct *Anguis* species (Grillitsch and Cabela, 1990) and the only species known to live in partial sympatry with another conspecific, *A. graeca*. It seems that the two species do not hybridize, or only rarely (Thanou et al., 2021). This is consistent with the nuclear genome topology and depths of divergences. These support a division into two main evolutionary lineages: *A. cephallonica* and a clade that we propose to systematically assess as the *A. fragilis* species complex, containing *A. fragilis*, *A. veronensis*, *A. colchica*, and *A. graeca*. The *A. fragilis* species complex began to diversify around 7 Mya. These four species forming the sister pairs *A. fragilis*-*A. veronensis* and *A. colchica*-*A. graeca* presently live in parapatry (Jablonski et al., 2021a) and hybridize in secondary contact zones formed between *A. fragilis* and *A. veronensis* (Gvoždík et al., 2013), *A. fragilis* and *A. colchica* (Szabó and Vörös, 2014; Benkovský et al., 2021), and *A. fragilis* and *A. graeca* (Mikulíček et al., 2018).

3.2.3. *Balkan-Apennine mitochondrial capture*

The onset of the *A. veronensis* mitonuclear discordance can be dated to around 5 Mya, when the split of *A. veronensis* mtDNA from *A. cephallonica* mtDNA was estimated to occur (Fig. 2). The dating of this split in mtDNA follows the Messinian event (Messinian salinity crisis), a geological event during which the Mediterranean Sea was partially or almost completely desiccated, beginning at approximately 5.97–5.96 Mya and ending with the Zanclean flood at 5.33 Mya (e.g., Periáñez and Abril, 2015; Mascle and Mascle, 2019; Andreetto et al., 2021). Although various authors have no consensus on the extent of the terrestrial land mass after the Mediterranean Sea drying event, there is general agreement that there was a land connection between the Apennine and Balkan Peninsulas, or that only a narrow water corridor persisted (e.g., Loget et al., 2006; Anzidei et al., 2014; Periáñez and Abril, 2015; Mascle and Mascle, 2019; Andreetto et al., 2021). We assume that the mitochondrial capture likely occurred during the Messinian event, upon a secondary contact of the ancestors of *A. fragilis-veronensis* and *A. cephallonica*, presumably within the present-day Apennine Peninsula or nearby and at a time when the *A. cephallonica* ancestor presumably had a wider distribution than the present-day *A. cephallonica* (Fig. 3A). After the re-separation of the present-day Apennine and Balkan Peninsulas, when the Zanclean flood refilled the Mediterranean basin, the geographically localized introgressed mtDNA (within the Apennines) began its own evolutionary history. We assume that this introgressed mtDNA was distributed only in a geographically restricted marginal population of the *A. fragilis-veronensis* ancestor, within the Apennine Peninsula (Fig. 3B). This marginal population gave rise to *A. veronensis* later during the Pliocene, when the water mass spread further around the Apennines (Pinna, 1989; Fig. 3C).

3.2.4. *Anguis colchica and A. graeca*

The speciation event separating the Balkan endemic *A. graeca* and Eastern European *A. colchica* (estimated at 5–4 Mya) may also have been triggered by the re-flooding of the Mediterranean basin at the onset of the Pliocene, which strongly affected local ecosystems and biota. The widespread species *A. colchica* began to diversify around the end of the Pliocene, when southeastern populations from the Ponto-Caspian region (including the Caucasus, where the species type locality is situated; *A. colchica* South-East clade, Fig. 2) separated from more northerly distributed populations corresponding to the currently recognized taxon *A. colchica incerta* (around 3 Mya). The *A. colchica* South-East clade further diversified during the Early Pleistocene at about 2.4 to 1.4 Mya. Three lineages are presently known from the region, two of them classified at the subspecific level (Gvoždík et al., 2010; Jablonski et al., 2016). However, a proper assessment of the subspecific taxonomy requires further investigation of the evolutionary history and geographic variation of the *A. colchica* South-East clade using denser sampling. This also applies to the evaluation of the possible species-level distinction of the relatively divergent *A. colchica incerta*, which occurs in allopatry (or parapatry in the Balkans) from the South-East clade of *A. colchica*.

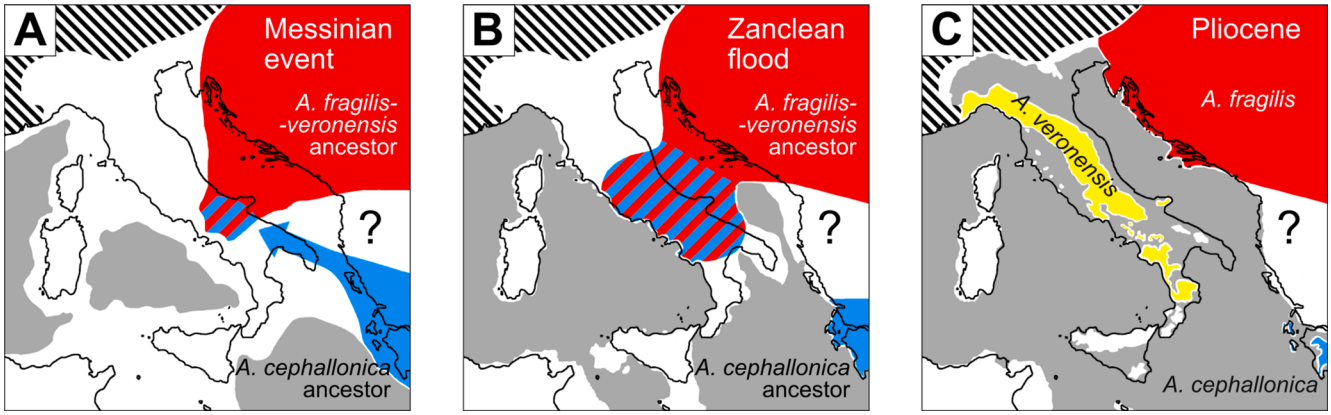

Fig. 3. Paleogeographic scenarios of the evolutionary history of *A. cephallonica*, *A. fragilis* and *A. veronensis* in the Balkan and Apennine Peninsulas. (A) Scenario of the mitochondrial capture from the Balkan ancestor of *A. cephallonica* ('proto-*A. cephallonica*', cyan) into a geographically restricted, peripheral, presumably Apennine population of the ancestor of *A. fragilis-veronensis* ('proto-*A. veronensis*, red-blue hatched) during the Messinian event. (B) Scenario during the Zanclean flood that ended the Messinian event, when 'proto-*A. veronensis*' gradually became geographically isolated. (C) Situation in the late Pliocene, when *A. veronensis* (yellow) speciated in allopatry. The distribution of water (grey) and land masses (white) was taken from various sources, mainly Pinna (1989), Loget et al. (2006), Periáñez and Abril (2015), Andreetto et al. (2021); the Alps are in black and white hatched area; the question mark indicates uncertainty about the extent of the presumed ancestral ranges of the discussed lineages. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2.5. Pseudopus

The glass lizard *Pseudopus apodus* diversified into three evolutionary lineages during the Pleistocene about 1.2 to 0.5 Mya. Thus, the extant *Pseudopus* representatives have only a shallow evolutionary history compared to the extant *Anguis*, which is consistent with some earlier studies (Jandzik et al., 2018; Jablonski et al., 2021a). The systematics of the extant taxa is therefore best described by taxonomy at the subspecies level.

4. Conclusions

Extant taxa of *Pseudopus* have shallow evolutionary history (Pleistocene) compared to extant taxa of *Anguis* (Late Miocene–Pliocene). The oldest division in *Anguis*, based on phylogenomic reconstruction of nuclear DNA, separates the southern Balkan endemic *A. cephallonica* from the widespread Western Palearctic *A. fragilis* species complex (*A. fragilis*, *A. veronensis*, *A. colchica*, *A. graeca*). The mitonuclear discordance in the placement of *A. veronensis* from the Apennine Peninsula is best explained by Balkan-Apennine mitochondrial capture during the Messinian salinity crisis. We hypothesize that the capture occurred only in a geographically restricted peripheral population in the ancestor of *A. fragilis-veronensis*, which later gave rise to *A. veronensis*. The taxonomy of *A. colchica* may be re-evaluated in the future considering the deep (late Pliocene) divergence separating the South-East (Ponto-Caspian) clade from the more northerly distributed taxon *A. colchica incerta*, and the latter may merit the species-level status. Finally, it is worth mentioning that despite the relatively high morphological similarity and geographic proximity, the studied individuals of *A. fragilis* and *A. colchica* from the vicinity of their hybrid zone in the Czech Republic were confirmed by the genomic analyses as deeply divergent (about 7 My), non-sister species.

Data availability

Sequence capture data obtained by the anchored hybrid enrichment (alignments of 383 loci) are available in the Mendeley Data repository: https://dx.doi.org/10.17632/37wvmkhxv4.1. New mitochondrial sequences were deposited in GenBank under the accession numbers OP493517–OP493530.

CRediT authorship contribution statement

Václav Gvoždík: Conceptualization, Methodology, Supervision, Funding acquisition, Resources, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Tadeáš Nečas:** Methodology, Investigation, Formal analysis, Visualization, Writing – review & editing. **Daniel Jablonski:** Resources, Investigation, Visualization, Writing – review & editing. **Emily Moriarty Lemmon:** Methodology, Writing – review & editing. **Alan R. Lemmon:** Methodology, Writing – review & editing. **David Jandzik:** Resources, Investigation, Writing – review & editing. **Jiří Moravec:** Funding acquisition, Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The statement is included in the manuscript.

Acknowledgements

We would like to express our gratitude to L. Choleva, M. Homolka, S. Meiri, R. Musilová, A. Romano, J. Šmíd, C. Spilinga and B.L. Stuart for their kind donation of samples, J. Klembara for helpful advice on the age of the oldest fossil records of *Anguis* and *Pseudopus*, and A. Hánová for her help in the laboratory. This study was funded by the Czech Science Foundation (grant number 18-24544S) and the Ministry of Culture of the Czech Republic (DKRVO 2019–2023/6.V.d, National Museum, 00023272). The work of D. Jablonski was supported by the Slovak Research and Development Agency under the contract APVV-19-0076 and T. Nečas by the Department of Botany and Zoology, Faculty of Science, Masaryk University (MUNI/A/1401/2021).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2022.107674.

References

Andreetto, F., Aloisi, G., Raad, F., Heida, H., Flecker, R., Agiadi, K., Lofi, J., Blondel, S., Bulian, F., Camerlenghi, A., Caruso, A., Ebner, R., Garcia-Castellanos, D., Gaullier, V., Guibourdenche, L., Gvirtzman, Z., Hoyle, T.M., Meijer, P.T., Moneron, J., Sierro, F.J., Travan, G., Tzevahirtzian, A., Vasiliev, I., Krijgsman, W., 2021. Freshening of the Mediterranean Salt Giant: controversies and certainties around the terminal (Upper Gypsum and Lago-Mare) phases of the Messinian Salinity Crisis. Earth-Sci. Rev. 216, 103577.

Anzidei, M., Lambeck, K., Antonioli, F., Furlani, S., Mastronuzzi, G., Serpelloni, E., Vannucci, G., 2014. Coastal structure, sea-level changes and vertical motion of the land in the Mediterranean, in: Martini, I.P., Wanless, H.R. (Eds.), Sedimentary Coastal Zones from High to Low Latitudes: Similarities and Differences. Geological Society, London, Special Publications 388, pp. 453–479.

Benkovský, N., Moravec, J., Gvoždíková Javůrková, V., Šifrová, H., Gvoždík, V., Jandzik, D., 2021. Phenotypic differentiation of the slow worm lizards (Squamata: *Anguis*) across their contact zone in Central Europe. PeerJ 9, e12482.

Bergsten, J., 2005. A review of long-branch attraction. Cladistics 21, 163–193.

Bouckaert, R., Heled, J., 2014. DensiTree 2: Seeing trees through the forest. bioRxiv, 012401.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Müller, N.F., Ogilvie, H.A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.-H., Xie, D., Zhang, C., Stadler, T., Drummond, A.J., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15, e1006650.

Burbrink, F.T., Grazziotin, F.G., Pyron, R.A., Cundall, D., Donnellan, S., Irish, F., Keogh, J.S., Kraus, F., Murphy, R.W., Noonan, B., Raxworthy, C.J., Ruane, S., Lemmon, A.R., Moriarty Lemmon, E., Zaher, H., 2020. Interrogating genomic-scale data for Squamata (lizards, snakes, and amphisbaenians) shows no support for key traditional morphological relationships. Syst. Biol. 69, 502–520.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Čerňanský, A., Klembara, J., Müller, J., 2016. The new rare record of the late Oligocene lizards and amphisbaenians from Germany and its impact on our knowledge of the European terminal Palaeogene. Palaeobio. Palaeoenv. 96, 559–587.

Dely, O.G., 1981. *Anguis fragilis* Linnaeus 1758 – Blindschleiche, in: Böhme, W. (Ed.), Handbuch der Reptilien und Amphibien Europas, Band 1 Echsen (Sauria) I. Akademische Verlagsgesellschaft, Wiesbaden, pp. 241–258.

Douglas, J., Zhang, R., Bouckaert, R., 2021. Adaptive dating and fast proposals: Revisiting the phylogenetic relaxed clock model. PLoS Comput. Biol. 17, e1008322.

Douglas, J., Jiménez-Silva, C.L., Bouckaert, R., 2022. StarBeast3: Adaptive parallelized Bayesian inference under the multispecies coalescent. Syst. Biol. 71, 901–916.

Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27, 401–410.

Grillitsch, H., Cabela, A., 1990. Zum systematischen Status der Blindschleichen (Squamata: Anguidae) der Peloponnes und der südlichen Ionischen Inseln (Griechenland). Herpetozoa 2, 131–153.

Gvoždík, V., Jandzik, D., Lymberakis, P., Jablonski, D., Moravec, J., 2010. Slow worm, *Anguis fragilis* (Reptilia: Anguidae) as a species complex: Genetic structure reveals deep divergences. Mol. Phylogenet. Evol. 55, 460–472.

Gvoždík, V., Benkovský, N., Crottini, A., Bellati, A., Moravec, J., Romano, A., Sacchi, R., Jandzik, D., 2013. An ancient lineage of slow worms, genus *Anguis* (Squamata: Anguidae), survived in the Italian Peninsula. Mol. Phylogenet. Evol. 69, 1077–1092.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

Jablonski, D., Jandzik, D., Mikulíček, P., Džukić, G., Ljubisavljević, K., Tzankov, N., Jelić, D., Thanou, E., Moravec, J., Gvoždík, V., 2016. Contrasting evolutionary histories of the legless lizards slow worms (*Anguis*) shaped by the topography of the Balkan Peninsula. BMC Evol. Biol. 16, 99.

Jablonski, D., Najbar, B., Grochowalska, R., Gvoždík, V., Strzała, T., 2017. Phylogeography and postglacial colonization of Central Europe by *Anguis fragilis* and *Anguis colchica*. Amphibia-Reptilia 38, 562–569.

Jablonski, D., Ribeiro-Júnior, M.A., Meiri, S., Maza, E., Kukushkin, O.V., Chirikova, M., Pirosová, A., Jelić, D., Mikulíček, P., Jandzik, D., 2021a. Morphological and genetic

differentiation in the anguid lizard *Pseudopus apodus* supports the existence of an endemic subspecies in the Levant. Vertebr. Zool. 71, 175–200.
Jablonski, D., Sillero, N., Oskyrko, O., Bellati, A., Čeirāns, A., Cheylan, M., Cogălniceanu, D., Crnobrnja-Isailović, J., Crochet, P.-A., Crottini, A., Doronin, I., Džukić, G., Geniez, P., Ilgaz, Ç., Iosif, R., Jandzik, D., Jelić, D., Litvinchuk, S., Ljubisavljević, K., Lymberakis, P., Mikulíček, P., Mizsei, E., Moravec, J., Najbar, B., Pabijan, M., Pupins, M., Sourrouille, P., Strachinis, I., Szabolcs, M., Thanou, E., Tzoras, E., Vergilov, V., Vörös, J., Gvoždík, V., 2021b. The distribution and biogeography of slow worms (*Anguis*, Squamata) across the Western Palearctic, with an emphasis on secondary contact zones. Amphibia-Reptilia 42, 519–530.
Jandzik, D., Jablonski, D., Zinenko, O., Kukushkin, O.V., Moravec, J., Gvoždík, V., 2018. Pleistocene extinctions and recent expansions in an anguid lizard of the genus *Pseudopus*. Zool. Scr. 47, 21–32.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.
Kapli, P., Yang, Z., Telford, M.J., 2020. Phylogenetic tree building in the genomic age. Nat. Rev. Genet. 21, 428–444.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Klembara, J., Rummel, M., 2018. New material of *Ophisaurus*, *Anguis* and *Pseudopus* (Squamata, Anguidae, Anguinae) from the Miocene of the Czech Republic and Germany and systematic revision and palaeobiogeography of the Cenozoic Anguinae. Geol. Mag. 155, 20–44.
Klembara, J., Hain, M., Dobiašová, K., 2014. Comparative anatomy of the lower jaw and dentition of *Pseudopus apodus* and the interrelationships of species of subfamily Anguinae (Anguimorpha, Anguidae). Anat. Rec. 297, 516–544.
Klembara, J., Hain, M., Čerňanský, A., 2019. The first record of anguine lizards (Anguimorpha, Anguidae) from the early Miocene locality Ulm – Westtangente in Germany. Hist. Biol. 31, 1016–1027.
Lavin, B.R., Girman, D.J., 2019. Phylogenetic relationships and divergence dating in the Glass Lizards (Anguinae). Mol. Phylogenet. Evol. 133, 128–140.
Lemmon, A.R., Emme, S.A., Lemmon, E.M., 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst. Biol. 61, 727–744.
Loget, N., Davy, P., Van Den Driessche, J., 2006. Mesoscale fluvial erosion parameters deduced from modelling the Mediterranean sea level drop during the Messinian (late Miocene). J. Geophys. Res. 111, F03005.
Macey, R.J., Schulte II, J.A., Larson, A., Tuniyev, B.S., Orlov, N., Papenfuss, T.J., 1999. Molecular phylogenetics, tRNA evolution, and historical biogeography in anguid lizards and related taxonomic families. Mol. Phylogenet. Evol. 12, 250–272.
Mascle, G., Mascle, J., 2019. The Messinian salinity legacy: 50 years later. Med. Geosc. Rev. 1, 5–15.
Mikulíček, P., Jablonski, D., Páleník, M., Gvoždík, V., Jandzik, D., 2018. Characterization of microsatellite markers in the genera *Anguis* and *Pseudopus* (Reptilia: Anguidae). Salamandra 54, 158–162.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534.
Obst, F.J., 1981. *Ophisaurus apodus* – Scheltopusik, Panzerschleiche, in: Böhme, W. (Ed.), Handbuch der Reptilien und Amphibien Europas, Band 1 Echsen (Sauria) I. Akademische Verlagsgesellschaft, Wiesbaden, pp. 259–274.
Periáñez, R., Abril, J.M., 2015. Computational fluid dynamics simulations of the Zanclean catastrophic flood of the Mediterranean (5.33 Ma). Palaeogeogr. Palaeoclimatol. Palaeoecol. 424, 49–60.
Pinna, G., 1989: Il periodo Neogenico, in: Pinna, G., Il grande libro dei fossili. Biblioteca Universale Rizzoli, pp. 253–257.
Pyron, R.A., Burbrink, F.T., Wiens, J.J., 2013. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evol. Biol. 13, 93.
Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Sabaj, M.H., 2022. Codes for Natural History Collections in Ichthyology and Herpetology (online supplement). Version 9.0 (14 February 2022). Electronically accessible at https://asih.org, American Society of Ichthyologists and Herpetologists, Washington, DC.
Szabó, K., Vörös, J., 2014. Distribution and hybridization of *Anguis fragilis* and *A. colchica* in Hungary. Amphibia-Reptilia 35, 135–140.
Thanou, E., Giokas, S., Kornilios, P., 2014. Phylogeography and genetic structure of the slow worms *Anguis cephallonica* and *Anguis graeca* (Squamata: Anguidae) from the southern Balkan Peninsula. Amphibia-Reptilia 35, 263–269.
Thanou, E., Kypraios-Skrekas, V., Kornilios, P., Giokas, S., 2021. Ecomorphological divergence and lack of gene flow in two sympatric Balkan slow worms (Squamata: Anguidae). Biol. J. Linn. Soc. 134, 443–460.
Villa, A., Delfino, M., 2019. Fossil lizards and worm lizards (Reptilia, Squamata) from the Neogene and Quaternary of Europe: an overview. Swiss J. Palaeontol. 138, 177–211.
Völkl, W., Alfermann, D., 2007. Die Blindschleiche – die vergessene Echse. Z. Feldherpetol., Beiheft 11. Laurenti-Verlag, Bielefeld.
Zhang, R., Drummond, A., 2020. Improving the performance of Bayesian phylogenetic inference under relaxed clock models. BMC Evol. Biol. 20, 54.
Zheng, Y., Peng, R., Kuro-o, M., Zeng, X., 2011. Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: A case study of salamanders (order Caudata). Mol. Biol. Evol. 28, 2521–2535.

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Supplementary material

Phylogenomics of *Anguis* and *Pseudopus* (Squamata, Anguidae) indicates Balkan-Apennine mitochondrial capture associated with the Messinian event

Václav Gvoždík*, Tadeáš Nečas, Daniel Jablonski, Emily Moriarty Lemmon, Alan R. Lemmon, David Jandzik, Jiří Moravec

* Corresponding author at:

Institute of Vertebrate Biology of the Czech Academy of Sciences, Research Facility Studenec, Studenec 122, 675 02, Czech Republic.

E-mail address: vaclav.gvozdik@gmail.com (V. Gvoždík).

CONTENT **Page**

Supplementary methods

Anchored hybrid enrichment

Illumina libraries were prepared from DNA extracts following the protocol of Meyer and Kircher (2010) with modifications outlined in Prum et al. (2015) at the Florida State University's Center for Anchored Phylogenomics (www.anchoredphylogeny.com). After adding 8 bp adapter indexes, libraries were pooled in equal concentrations, then enriched using the lizard-specific probe set described in Tucker et al. (2016). Enriched libraries were sequenced on one Illumina HiSeq2500 lane with a PE150bp protocol in the Translational Lab at the FSU College of Medicine. The total sequence data collected was 11.2 Gb.

After demultiplexing the Illumina reads, adapters were removed, and sequencing errors were corrected while overlapping read pairs were merged as outlined in Rokyta et al. (2012). Reads were assembled into clusters using a quasi-denovo assembly approach described by Hamilton et al. (2016). Consensus sequences were constructed from assembly clusters containing at least 117 reads, a threshold which allowed the removal of low-copy paralogs and any potential low-level contamination, while still retaining the large majority of target loci. After homologous consensus sequences were gathered across the 14 individuals, orthology was assessed using a neighbor-joining approach that utilized an alignment-free pairwise distance matrix (methodological details provided in Hamilton et al., 2016). Orthologous sets of sequences were then aligned using MAFFT v7 (Katoh and Standley, 2013). Raw alignments were subjected to an automated trimmer/masker that identifies misaligned regions, as described in Hamilton et al. (2016; but with parameters: MINGOODSITES=16, MINPROPSAME=0.3). We removed sites represented by fewer than 7 of the 14 terminals to mitigate the possible effects of missing data (Lemmon et al., 2009). The alignments were inspected by eye in Geneious R9 (https://www.geneious.com) to ensure the automated trimmer/masker settings were suitable.

References

Hamilton, C.A., Lemmon, A.R., Lemmon, E.M., Bond, J.E., 2016. Expanding anchored hybrid enrichment to resolve both deep and shallow relationships within the spider tree of life. BMC Evol. Biol. 16, 212.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Lemmon, A.R., Brown, J.M., Stanger-Hall, K., Lemmon, E.M., 2009. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Syst. Biol. 58, 130–145.

Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. 2010, 5448. Doi:10.1101/pdb.prot5448

Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M., Lemmon, A.R., 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526, 569–573.

Rokyta, D.R., Lemmon, A.R., Margres, M.J., Aronow, K., 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). BMC Genomics 13, 312.

Tucker, D.B., Colli, G.R., Giugliano, L.G., Hedges, S.B., Hendry, C.R., Lemmon, E.M., Lemmon, A.R., Sites Jr., J.W., Pyron, R.A., 2016. Methodological congruence in phylogenomic analyses with morphological support for teiid lizards (Sauria: Teiidae). Mol. Phylogenet. Evol. 103, 75–84.

Table S1. Partitions and substitution models used in **nDNA Bayesian analyses** (MrBayes, StarBeast3).

Partition	Composition (L = locus)	Model
1	L1; L17; L20; L25; L77; L96; L119; L223; L229; L234; L288; L292; L295; L334; L353; L386	HKY+I
2	L2; L176; L235; L329	HKY+I
3	L3; L55; L82; L97; L217; L294; L318; L342	GTR+I
4	L4; L178; L187; L188; L208; L265; L272; L287; L326; L387	HKY+I
5	L5; L6; L18; L19; L29; L56; L84; L138; L150; L193; L207; L219; L232; L238; L251; L252; L258; L264; L270; L285; L290; L293; L301; L303; L305; L315; L316; L323; L332; L335; L338; L376	HKY+I
6	L7; L36; L46; L47; L51; L76; L114; L125; L148; L149; L174; L184; L186; L191; L198; L201; L204; L206; L215; L246; L247; L248; L253; L263; L276; L300; L302; L336; L344; L352; L356; L371	HKY+I
7	L8; L78; L101; L102; L231; L237; L262; L321	HKY+G
8	L9; L92; L129; L130; L153; L200; L243; L312; L345; L348; L349; L351; L358; L361; L365; L384	HKY+I
9	L10; L23; L24; L33; L40; L80; L93; L108; L115; L143; L185; L190; L214; L216; L224; L225; L255; L256; L257; L259; L260; L268; L279; L280; L296; L313; L337; L350; L355; L367; L368; L375	HKY+I
10	L11; L13; L28; L31; L32; L41; L58; L66; L68; L83; L89; L98; L104; L110; L121; L124; L140; L160; L169; L173; L202; L211; L212; L244; L245; L275; L281; L298; L325; L369	HKY+I
11	L12; L30; L48; L69; L71; L73; L74; L88; L109; L111; L126; L132; L135; L136; L139; L147; L159; L165; L181; L182; L233; L242; L277; L282; L284; L286; L289; L310; L347; L360; L379; L381	HKY+I
12	L14; L43; L53; L54; L79; L113; L118; L152; L271; L274; L283; L311; L314; L320; L328; L330	HKY+I
13	L15; L16; L38; L49; L62; L85; L103; L105; L131; L154; L163; L166; L167; L179; L180; L189; L210; L221; L240; L299; L307; L324; L346; L362	HKY+I
14	L21; L37; L44; L57; L59; L65; L70; L75; L86; L106; L133; L142; L144; L146; L155; L156; L158; L161; L170; L175; L199; L213; L220; L306; L308; L317; L340; L354; L363; L370; L372; L377	HKY+I
15	L26; L45; L123; L134; L145; L205; L218; L227; L261; L267; L269; L319; L366; L374; L383; L385	HKY+I
16	L27; L39; L203; L341	HKY+I
17	L34; L35; L99; L100; L196; L322; L327; L357	GTR+I
18	L42; L137; L141; L164; L177; L194; L197; L226; L291; L333; L378; L380	HKY+I
19	L50; L52; L63; L72; L87; L90; L94; L107; L112; L116; L117; L120; L127; L128; L157; L162; L171; L183; L192; L209; L236; L241; L266; L273; L297; L304; L309; L343; L359	HKY+I
20	L61; L64; L95; L122; L168; L222; L228; L230; L239; L254; L278; L331; L339; L364; L373; L382	HKY+I
21	L81; L91; L151; L172	HKY+I
22	L195	GTR+I

Table S2. Partitions and substitution models used in the **nDNA maximum likelihood** (IQ-TREE) analysis.

Partition	Composition (L = locus)	Model
1	L1; L17; L20; L25; L77; L96; L119; L223; L229; L234; L288; L292; L295; L334; L353; L386	K3Pu+I
2	L2; L176; L235; L329	HKY+I
3	L3; L55; L82; L97; L217; L294; L318; L342	TN+I
4	L4; L178; L187; L188; L208; L265; L272; L287; L326; L387	HKY+I
5	L5; L18; L19; L29; L84; L138; L150; L238; L258; L264; L285; L303; L315; L316; L338; L376	K3Pu+I
6	L6; L10; L16; L46; L49; L57; L75; L167; L184; L193; L201; L207; L219; L232; L239; L240; L248; L251; L252; L289; L290; L293; L296; L301; L305; L307; L323; L324; L332; L335; L340; L372	TPM3+I
7	L7; L47; L51; L63; L76; L90; L107; L114; L128; L148; L149; L162; L171; L186; L191; L192; L204; L206; L209; L236; L246; L253; L273; L297; L309; L336; L343; L344; L359; L371	TIM3+I
8	L8; L78; L101; L102; L231; L237; L262; L321	TIM+G
9	L9; L92; L129; L130; L153; L200; L243; L312; L345; L348; L349; L351; L358; L361; L365; L384	HKY+I
10	L11; L13; L31; L32; L41; L58; L66; L68; L83; L89; L98; L104; L110; L121; L124; L140; L160; L169; L202; L211; L212; L230; L242; L244; L245; L275; L281; L298; L325; L339; L364; L369	TN+I
11	L12; L28; L30; L48; L69; L71; L73; L74; L88; L109; L111; L125; L126; L135; L136; L139; L159; L173; L181; L182; L198; L233; L277; L282; L284; L286; L300; L347; L356; L360; L379; L381	K3Pu+I
12	L14; L24; L33; L43; L53; L54; L79; L80; L93; L113; L118; L143; L152; L260; L271; L274; L283; L311; L314; L320; L328; L330; L355; L367	TN+R3
13	L15; L38; L56; L61; L62; L85; L103; L105; L131; L154; L163; L166; L179; L180; L189; L210; L221; L222; L270; L299; L331; L346; L362; L382	K3Pu+I
14	L21; L23; L37; L44; L59; L70; L106; L108; L115; L133; L142; L144; L146; L156; L158; L161; L175; L199; L213; L214; L216; L224; L225; L257; L268; L306; L308; L337; L350; L354; L370; L375	HKY+I
15	L26; L45; L123; L134; L145; L205; L218; L227; L261; L267; L269; L319; L366; L374; L383; L385	TPM2+I
16	L27; L39; L203; L341	TPM3+I
17	L34; L35; L40; L99; L100; L185; L190; L196; L255; L256; L259; L279; L280; L313; L322; L327; L357; L368	TN+R2
18	L36; L42; L141; L174; L177; L194; L197; L215; L226; L247; L263; L276; L302; L333; L352; L378	HKY+I
19	L50; L52; L67; L72; L87; L94; L112; L116; L117; L120; L127; L132; L137; L147; L157; L164; L165; L183; L241; L266; L291; L304; L310; L380	TPM2+I
20	L64; L65; L86; L95; L122; L155; L168; L170; L220; L228; L254; L278; L317; L363; L373; L377	HKY+I
21	L81; L91; L151; L172	TN+I
22	L195	TPM3+I

Table S3. Partitions and substitution models used in **mtDNA Bayesian analyses** (MrBayes, Standard BEAST).

Partition	Composition	Model
1	Phe; 12S; Val; 16S; ND1/1; Ile; Gln; Met; ND2/1; Trp; Ala; Cys; Tyr; Ser; Asp; COX2/1; Lys; ATP6/1; COX3/2; Gly; ND3/1; Arg; ND4L/1; ND4/1; His; Ser; Leu; ND5/1; Glu; CYTB/1; Thr; Pro	GTR+I+G
2	Leu; ND6/3	HKY+G
3	ND1/2; ND2/2; ATP6/2; ND3/2; ND4L/2; ND4/2; ND5/2; CYTB/2	HKY+I
4	ND1/3; ND2/3; COX1/1; COX2/3; ATP8/3; ATP6/3; COX3/3; ND3/3; ND4L/3; ND4/3; ND5/3; CYTB/3	GTR+I+G
5	Asn; COX1/2; COX1/3; COX2/2; ATP8/2; COX3/2	HKY+I+G
6	ATP8/1; CR	HKY+G
7	ND6/1; ND6/2	HKY+G

Table S4. Partitions and substitution models used in the **mtDNA maximum likelihood** (IQ-TREE) analysis.

Partition	Composition	Model
1	Phe; 12S; Val; 16S; ND1/1; Ile; Gln; Met; ND2/1; Trp; Ala; Cys; Tyr; Ser; Asp; COX2/1; Lys; ATP6/1; COX3/1; Gly; ND3/1; Arg; ND4L/1; ND4/1; His; Ser; Leu; ND5/1; Glu; CYTB/1; Thr; Pro	TIM2+R2
2	Leu; ND6/3	TN+I
3	ND1/2; ND2/2; ATP6/2; ND3/2; ND4L/2; ND4/2; ND5/2; CYTB/2	TN+I
4	ND1/3; ND2/3; COX1/1; COX2/3; ATP8/3; ATP6/3; COX3/3; ND3/3; ND4L/3; ND4/3; ND5/3; CYTB/3	TIM3+I+G
5	Asn; COX1/2; COX1/3; COX2/2; ATP8/2; COX3/2	TN+R2
6	ATP8/1; CR	TPM2+G
7	ND6/1; ND6/2	TN+G

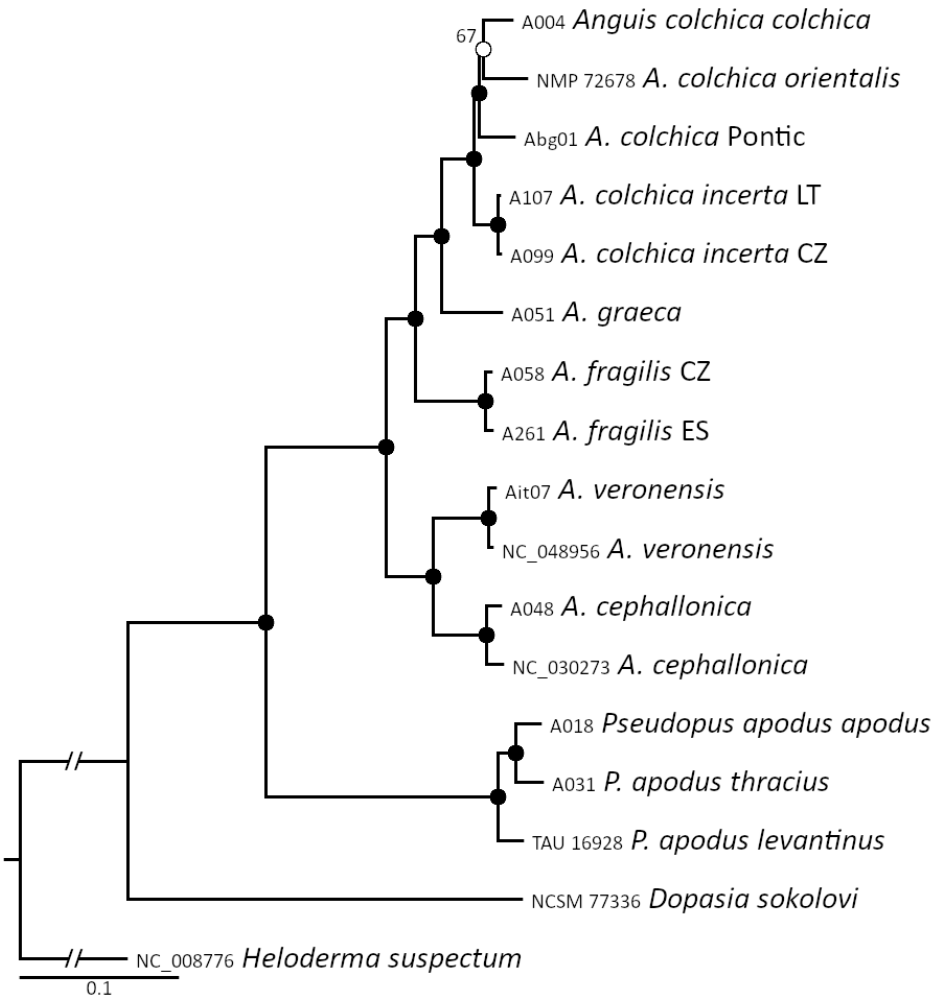

Fig. S1. Maximum likelihood (ML) mitogenomic tree with an additional outgroup, *Heloderma suspectum* (Helodermatidae). Black dot = bootstrap support > 70. Codes starting with "NC_" are sequences taken from GenBank.

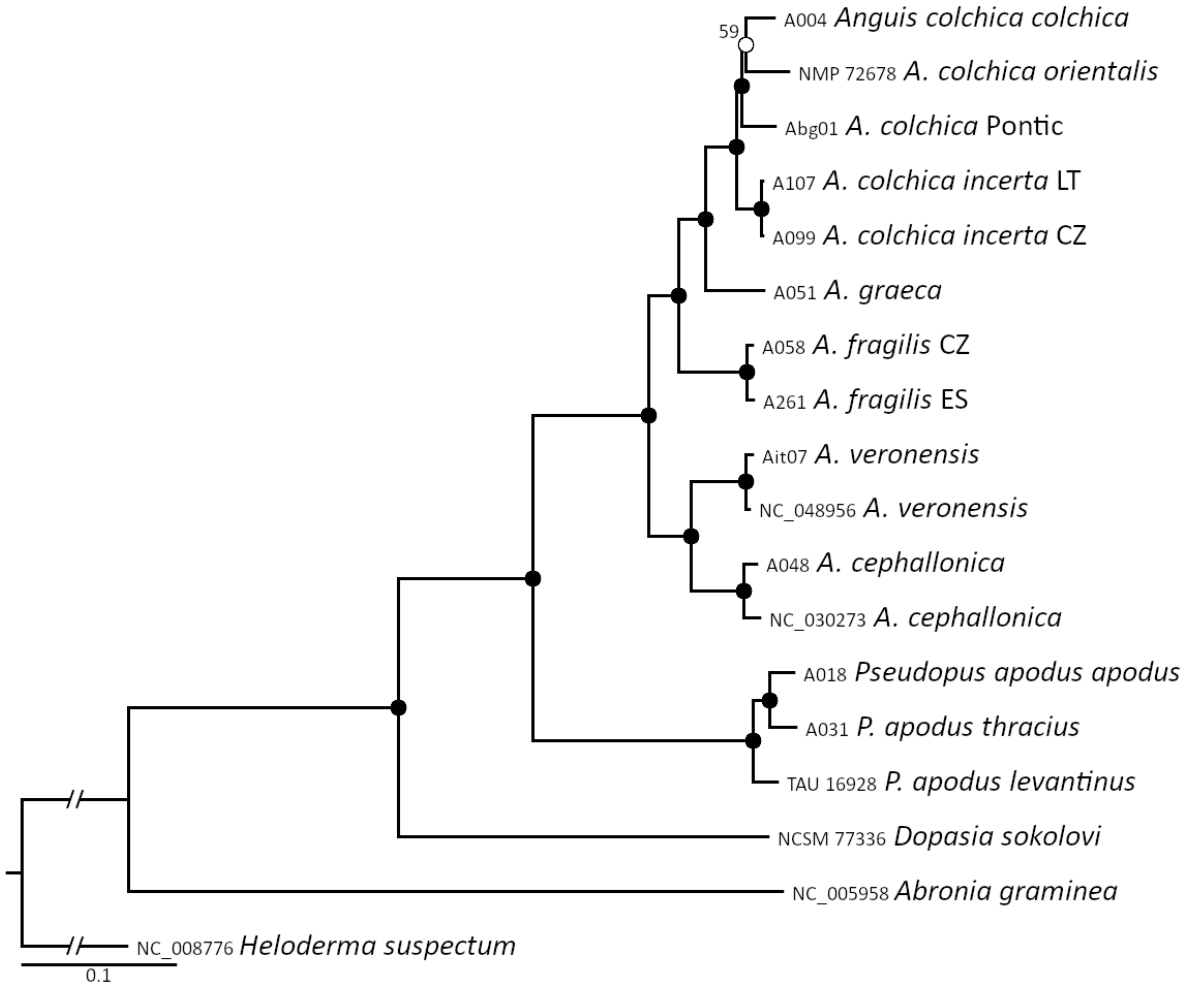

Fig. S2. ML mitogenomic tree with two additional outgroups, *Abronia graminea* (Anguidae, Gerrhonotinae) and *Heloderma suspectum* (Helodermatidae). Black dot = bootstrap support > 70. Codes starting with "NC_" are sequences taken from GenBank.

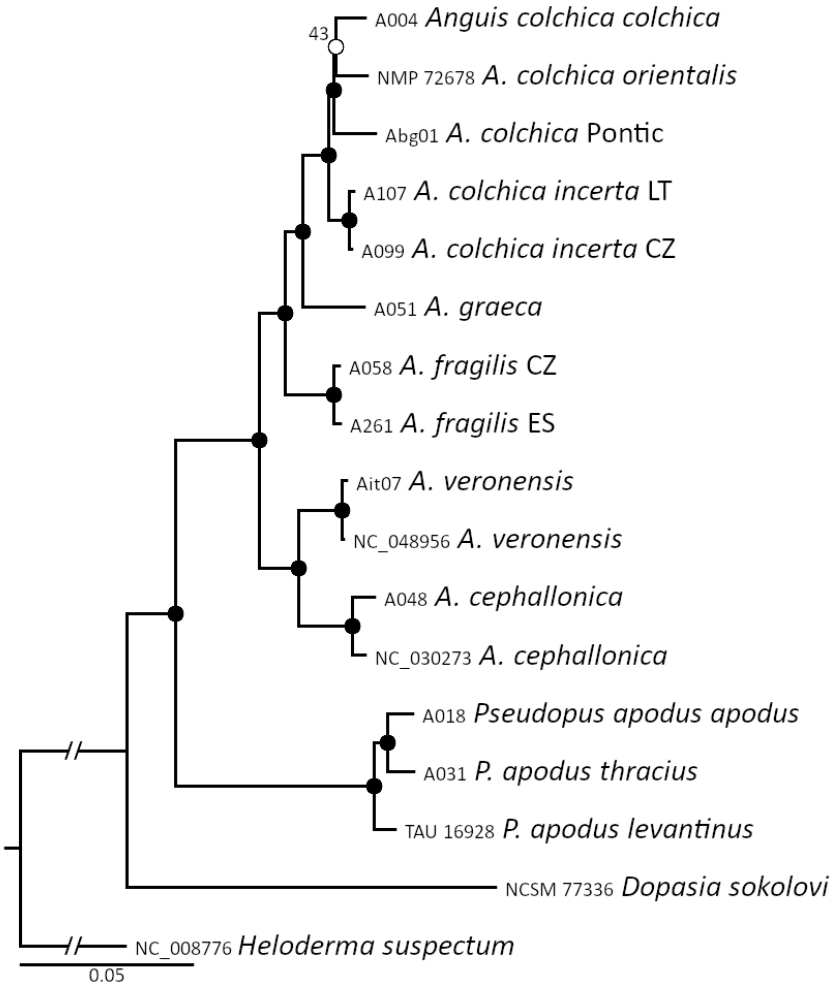

Fig. S3. ML mitogenomic tree after removing third codon positions of protein-coding loci and ambiguously aligned positions of non-protein-coding loci (alignment: 12,131 bp). Black dot = bootstrap support > 70. Codes starting with "NC_" are sequences taken from GenBank.

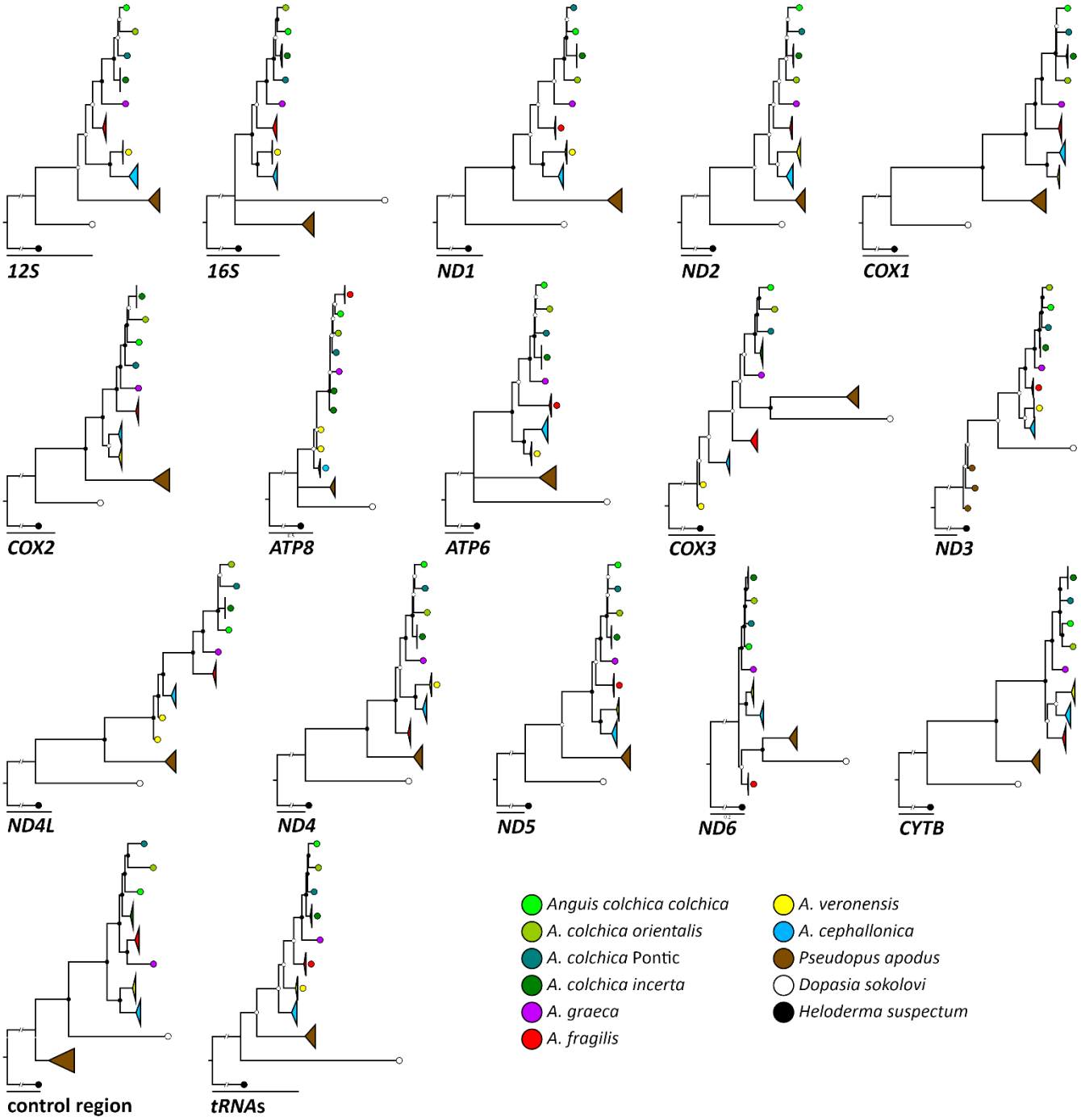

Fig. S4. ML trees of individual mtDNA loci (*tRNA*s considered as 'one locus'). Black dot in nodes = bootstrap support ≥ 70; white dot in nodes = bootstrap support < 70; scale = 0.10 substitution/site, except for *ATP8* (0.50 substitution/site) and *ND6* (0.20 substitution/site). Sequences of two additional *Anguis* individuals (*A. cephallonica*, *A. veronensis*) and *Heloderma* were taken from GenBank, see Fig. S1.