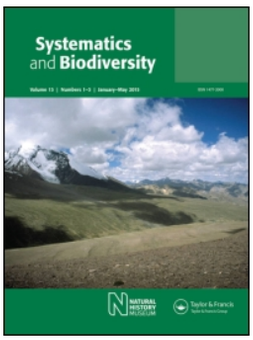

Systematics and Biodiversity

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsab20

Mitogenome evolution in the *Lacerta viridis* complex (Lacertidae, Squamata) reveals phylogeny of diverging clades

Robin-Tobias Jauss, Nadiné Solf, Sree Rohit Raj Kolora, Stefan Schaffer, Ronny Wolf, Klaus Henle, Uwe Fritz & Martin Schlegel

To cite this article: Robin-Tobias Jauss, Nadiné Solf, Sree Rohit Raj Kolora, Stefan Schaffer, Ronny Wolf, Klaus Henle, Uwe Fritz & Martin Schlegel (2021): Mitogenome evolution in the *Lacerta viridis* complex (Lacertidae, Squamata) reveals phylogeny of diverging clades, Systematics and Biodiversity, DOI: 10.1080/14772000.2021.1912205

To link to this article: https://doi.org/10.1080/14772000.2021.1912205

Systematics and Biodiversity (2021), **0(0)**: 1–12

Research Article

Mitogenome evolution in the *Lacerta viridis* complex (Lacertidae, Squamata) reveals phylogeny of diverging clades

ROBIN-TOBIAS JAUSS[1]*, NADINÉ SOLF[1]*, SREE ROHIT RAJ KOLORA[2], STEFAN SCHAFFER[3,4], RONNY WOLF[3], KLAUS HENLE[4,5], UWE FRITZ[6]* & MARTIN SCHLEGEL[1,4]*

[1]Institute of Biology, Biodiversity & Evolution, University of Leipzig, Talstraße 33, Leipzig, 04103, Germany
[2]Department of Integrative Biology, University of California Berkeley, Berkeley, CA, USA
[3]Institute of Biology, Molecular Evolution & Animal Systematics, University of Leipzig, Talstraße 33, Leipzig, 04103, Germany
[4]German Centre for Integrative Biodiversity Research (iDiv) Halle Jena Leipzig, Deutscher Platz 5e, Leipzig, 04103, Germany
[5]Department of Conservation Biology, UFZ – Helmholtz Centre for Environmental Research, Permoserstr. 15, 04318, Leipzig, Germany
[6]Museum of Zoology, Senckenberg Dresden, A. B. Meyer Building, 01109, Dresden, Germany

(Received 25 November 2020; accepted 29 March 2021)

In order to elucidate mitochondrial evolution and resolve the phylogeny of the widely distributed European green lizards (*Lacerta viridis* complex), we determined 16 mitochondrial genomes from a representative sampling of the four previously identified major clades corresponding to *L. bilineata* (western distribution range), *L. viridis* (most of the eastern distribution range), the unnamed Adriatic (or West Balkan) and the Turkish + south-eastern European clades. The mitogenomes were on average 17,162 bp long with a canonical vertebrate gene inventory of 13 protein-coding genes, 22 transfer RNA genes, 2 ribosomal RNA genes and a long non-coding Control Region. We detected several almost completely clade-specific insert and tandem-repeat signatures. Phylogenetic analyses using all 13 protein-coding genes as well as rRNA- and tRNA-coding sequences from a total of 20 mitogenomes inferred a resolved branching pattern of the four divergent clades. The Turkish clade is strongly supported as sister to all other members of the complex. Within the monophylum comprising the remaining clades, *L viridis* is sister to a clade containing *L. bilineata* and the Adriatic clade. Unexpectedly, one specimen of western Italy (Calambrone, Pisa) – which is within the distribution range of *L. bilineata* – clustered together with the Adriatic clade with high bootstrap support, calling for further research on the biogeography of the *L. viridis* complex. Integrating our results with those of previous studies suggest that the lineages have reached the level of distinct taxa, but to determine whether they have become fully independent lineages on the species level requires further research that tackles nuclear genomic variation of all four clades and the viability of hybrids and gene flow within putative hybrid zones.

Key words: biogeography, diversification, green lizard lineages, mitochondrial genes, phylogeny

Introduction

Background

Climatic changes had a strong impact on speciation and phylogeographic patterns in Europe (Hewitt, 2000; Taberlet et al., 1998). A prime example is the widely distributed European green lizard (*Lacerta viridis* complex, Lacertidae, Squamata), with currently two recognized species, one in the western (*L. bilineata*) and another in the eastern (*L. viridis*) part of the distribution range. A putative contact and hybrid zone was suggested for north-eastern Italy and western Slovenia (Joger et al., 2001; Nettmann, 2001). However, their separation in two distinct species has been a matter of an on-going controversy (Böhme et al., 2007b; Godinho et al., 2005; Marzahn et al., 2016). Expanded sampling and screening using mtDNA markers revealed that the putative hybrid zone between *L. bilineata* and *L. viridis* most likely does not exist because a third, equally divergent mitochondrial clade, the so-called West Balkan or Adriatic clade (Böhme et al., 2007b; Marzahn et al., 2016) geographically separates

Correspondence to: Martin Schlegel. E-mail: schlegel@uni-leipzig.de
*These authors contributed equally and should be considered joint first and last authors.

ISSN 1477-2000 print / 1478-0933 online

https://dx.doi.org/10.1080/14772000.2021.1912205
Published online 18 May 2021

the mitochondrial clades corresponding to *L. bilineata* and *L. viridis*. Moreover, a fourth divergent clade was discovered in northern Anatolia (Turkey) and adjacent Europe (Marzahn et al., 2016). The phylogenetic branching patterns of these four clades are weakly resolved. A previous analysis used the mitochondrial cytochrome *b* gene and found all four clades as an unresolved polytomy (Marzahn et al., 2016). Thus, further research is needed to tackle this issue. Moreover, it remains unclear whether these mitochondrial clades reflect distinct species or conspecific variation.

Mitochondrial genomes as phylogenetic tools

Mitochondrial genomes have been used as phylogenetic tools in various studies, including European green lizards (Bernt et al., 2013a; Böhme et al., 2007a; Kehlmaier et al., 2020; Kolora et al., 2017; Mendes et al., 2016; Perseke et al., 2013). Recent improvements in determining whole genomes raised the number of animal mitogenomes to some 10,000 (NCBI RefSeq). Their advantage in phylogenetic analyses is two-fold: firstly, the relatively high number of gene sequences, which evolve faster than nuclear markers and, as in our case, are more suitable to resolve recent diversification events, and secondly, as a structural character, the variation in gene orders (Bernt et al., 2007; 2013a; Fritzsch et al., 2006; Perseke et al., 2008). On the other hand, in the majority of animal species, mitochondria are inherited maternally, and thus harbour some disadvantages compared to nuclear markers, i.e. they do not yield information on gene flow between populations, such as in contact zones, and thus do not allow conclusions on the species status of even deep branching clades. Only tentative inferences can be made, for example by comparing genetic distances to closely related congeners that represent undoubtedly distinct species.

Aims of this study

Here, we determined the complete mitochondrial genomes of at least two representatives of each of the four divergent clades of the *L. viridis* complex with two aims in mind: (i) Do complete mitogenomes resolve robustly the phylogeny of the four clades? (ii) Does a comparison of genetic distances with congeners allow taxonomic inferences?

Materials and methods

Origin of samples

Samples originated from the tissue collections of the Institute of Biology of the University of Leipzig, the Museum of Zoology, Senckenberg Dresden, and the Natural History Museum in Vienna. In addition, we included four mitogenomes from GenBank in our analyses. Our material covers the whole diversity of mitochondrial variation of the *L. viridis* complex. Detailed sampling information is provided in Supplementary Table S1.

DNA isolation, library preparation and sequencing

Genomic DNA was isolated from either blood or tissue samples with DNA Blood and Tissue Kits (Qiagen) following the instructions of the manufacturer. Isolation was repeated with the DTAB method (Gustinich et al., 1991) to increase the yield, followed by a quality check and adjustment of DNA concentration using a NanoDrop UV-Vis Spectrophotometer (Thermo Fisher Scientific). Quality control was repeated for library preparation with a Qubit4 Fluorometer (Thermo Fisher Scientific). For preparation of paired-end libraries, DNA was sheared on a Covaris Sonicater to 300-350 bp fragment length, followed by blunt end repair, adapter ligation and fill-in with P5 and P7 adapters. Indexing was performed using the Nextera DNA CD Indexes Kit (Dabney & Meyer, 2012). After heteroduplex removal, samples were sequenced on an Illumina MiSeq platform with a read length of 150 bp (forward and reverse). Reads were processed with Trimmomatic version 0.39 (Bolger et al., 2014) using the following parameters: Illuminaclip 2:30:10, leading: 3, trailing: 3, minimum phred score of 30 in a 4 bp sliding window, minimum read length: 40 bp. Quality of reads was assessed with FastQC 0.11.8 (Andrews, 2010) before and after trimming.

Mitogenome assembly and annotation

Mitogenomes were assembled with MITObim version 1.9.1 (Hahn et al., 2013; Hahn & Schultz, 2018). Reads were baited with the mitogenome of *Lacerta viridis viridis,* GenBank accession number AM176577.1 (Böhme et al., 2007a). Trimmed reads were mapped against contigs using Bowtie2, version 2.3.5.1 (Langmead & Salzberg, 2012) and visualized with Tablet viewer (Milne et al., 2010) for inspection of coverage and assembly errors. The assembled mitogenomes were annotated using MITOS2 (Bernt et al., 2013b) and compared with the available outgroup mitogenomes of *L.*

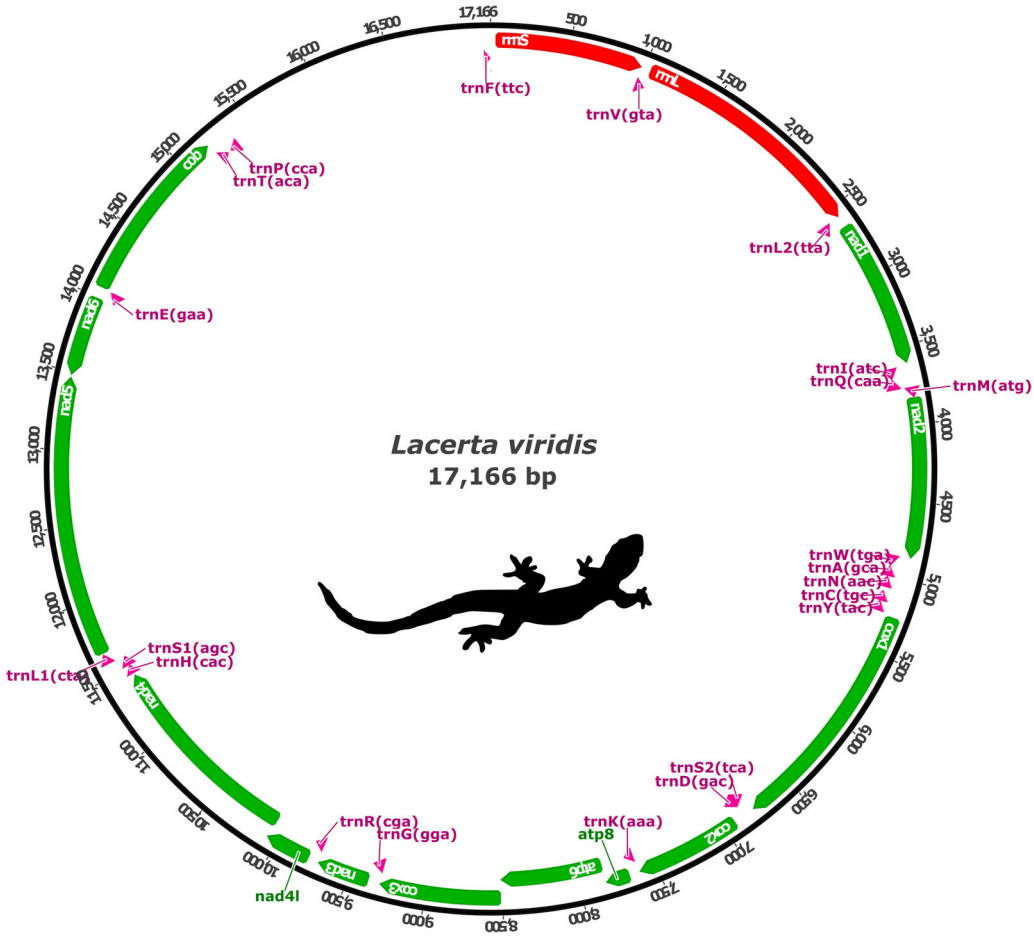

Fig. 1. Mitochondrial genome of *Lacerta viridis* G664_Lv_CZ. Green: protein-coding genes; red: ribosomal genes; pink: transfer RNA genes. The Control Region is located between trnP and trnF. Nucleotide positions are indicated clockwise.

agilis NC_021766.1 (Qin et al., 2014), *Zootoca vivipara* NC_990830.1 (Liu et al., 2016), *Podarcis muralis* NC_011607.1 (Podnar et al., 2007), *Takydromus wolteri* NC_018777.1 (Yu & Ji, 2013), and *Eremias vermiculata* NC_025320.1 (Tong et al., 2016). Assembled genomes were visualized and pairwise distances between genomes were calculated with Geneious version 9.0.5 (https://www.geneious.com).

Phylogenetic analyses

Extraction and alignment of protein-coding genes were performed using custom BioRuby-Scripts, version 2.0.0 (Goto et al., 2010). Nucleotide sequences were translated into amino acid sequences in MEGA X (Kumar et al., 2018) using the vertebrate mitochondrial genetic code. Single genes were aligned with MAFFT, version 7.428 with the setting "local pair" (Katoh & Standley, 2013) and visually inspected in MEGA X. Concatenation of nucleotide and amino acid alignments were performed with BioRuby-Scripts. The most suitable evolutionary models for proteins were determined with ProtTest 3, version 3.4.2 (Darriba et al., 2011), nucleotide models were determined with jModelTest version 2.1.10 (Darriba et al., 2012). The best-fitting models were selected using the Akaike information criterion (Akaike, 1974).

For phylogenetic analyses three published mitogenomes were included: *L. viridis* (AM176577.1), Adriatic clade (ERS4631750), *L. bilineata* (KT722705.1) and another *L. viridis* mitogenome determined by Kolora et al. (2017) but submitted to ENA within this study (ERZ1645122, Table S1). For a comprehensive inference, phylogenetic trees were based on nucleotide sequences of all 13 protein-coding genes concatenated with all rRNA- and tRNA-coding sequences. Additionally, amino acid sequences of the 13 protein-coding genes were used. Trees were reconstructed with a Maximum Likelihood approach using RAxML 8.2.12 (Stamatakis, 2006) and the GTR + G evolutionary model. Support for clades was quantified with bootstrap analyses of 1,000 pseudoreplicates (Felsenstein, 1985).

Fig. 2. Specific DNA sequence motifs within the Control Region (CR). Motifs include two flanking t-RNA genes, three Conserved Sequence Blocks, one 11 bp Insert, Imperfect and Perfect Repeats as well as clade-specific Termination Associated Sequences.

To test alternative topologies, additional constrained trees enforcing the monophyly of *L. viridis* samples + Turkish clade samples as well as *L. viridis* samples + Adriatic clade samples, respectively, were reconstructed with RAxML and evaluated with an approximately unbiased (AU) test (Shimodaira, 2002) as implemented in IQ-TREE version 1.6.12 (Nguyen et al., 2015). Bayesian Inference trees were calculated using MrBayes 3.2.6 (Altekar et al., 2004; Ronquist et al., 2012) and the GTR + G model. The number of generations was set to 1,000,000 with 4 Markov chains. All trees were processed in MEGA X and Paint NET 4.2 and visualized using R 3.6 (R Core Team, 2019) with the packages phytools 0.7-20 (Revell Liam, 2012) and ggplot2 3.3 (Wickham, 2016). Gene saturation tests were performed by plotting the number of transitions and transversions against the raw genetic distance for each protein-coding gene, as implemented in the R package ape version 5.4-1 (Paradis & Schliep, 2019).

Results

Mitogenomes

We obtained 15 complete and one near-complete mitochondrial genomes, the latter lacking a 23-bp fragment within the Control Region. The median coverage was 16x per nucleotide position (Supplementary Table S2).

The median length of the mitogenomes was 17,162 bp (Supplementary Table S2), with a canonical vertebrate inventory of 13 protein-coding genes, 22 tRNAs, 2 rRNAs, and a long non-coding Control Region (CR) (Fig. 1). Here, however, the positions of genes coding for tRNA-L1 and tRNA-L2 were switched compared to the mitogenome of a *Lacerta viridis viridis,* accession number AM176577.1 (Böhme et al., 2007a). Nucleotide composition was 28% T, 27% C, 31.4% A, 13.2% G (40.2% GC).

Clade-specific signatures in the control region

The median length of the Control Region (CR) located between tRNA-P and tRNA-F was 1823 bp. The CR contained specific sequence motifs (Fig. 2). The 5′-end comprised a 35 bp-sequence motif in eight repeats, except for one Adriatic lizard (G2701_AL_SI) and one *L. bilineata* (KT722705.1) with six repeats. This repeat region showed an increase in C (36.1%) and a decrease of A (16.9%) in contrast to the whole CR region (22.7% C, 30.5% A). Two Termination Associated Sequences (TAS-1, TAS-2) were located further downstream. Three almost clade-specific variants of TAS-2 were found (TAS-2a, TAS-2b, TAS-2c; Fig. 2, Supplementary Table S3). Adjacent to TAS-1 and TAS-2 a completely conserved 14 bp sequence repeat was positioned. Two incomplete repeats with some variation between the samples were located in the centre of the CR, followed further downstream by three identically conserved sequence blocks (CSB 1-3): All *L. viridis* had an 11 bp insertion downstream of CSB-1 that was missing in all other samples (Fig. 2, Supplementary Table S3).

Phylogenetic analyses

Gene saturation tests revealed little to no saturation for the protein-coding genes, indicating a robust dataset for accurate phylogenetic reconstructions (Fig. S1). The phylogenetic analyses using Maximum Likelihood (ML) and Bayesian Inference methods for the nucleotide sequences of the 13 protein-coding genes plus the sequences coding for rRNA and tRNA yielded almost the same and well-supported tree topology compared to that using the amino acid sequences of the 13 protein-coding genes. Therefore, we restrict our results and the discussion to the results based on the concatenated sequences of the 13 protein-coding genes plus those for rRNA and tRNA (Figs. 3–4; S2-S3).

The monophyly of all sampled individuals of the *L. viridis* complex received maximum support. Within the complex, the Turkish clade was sister to all other samples. The monophyly of the remaining clades was well supported (95% bootstrap support). Within this monophylum, all samples of *L. viridis* constituted a clade with maximum support that was sister to a clade consisting of *L. bilineata* and the Adriatic clade, which was

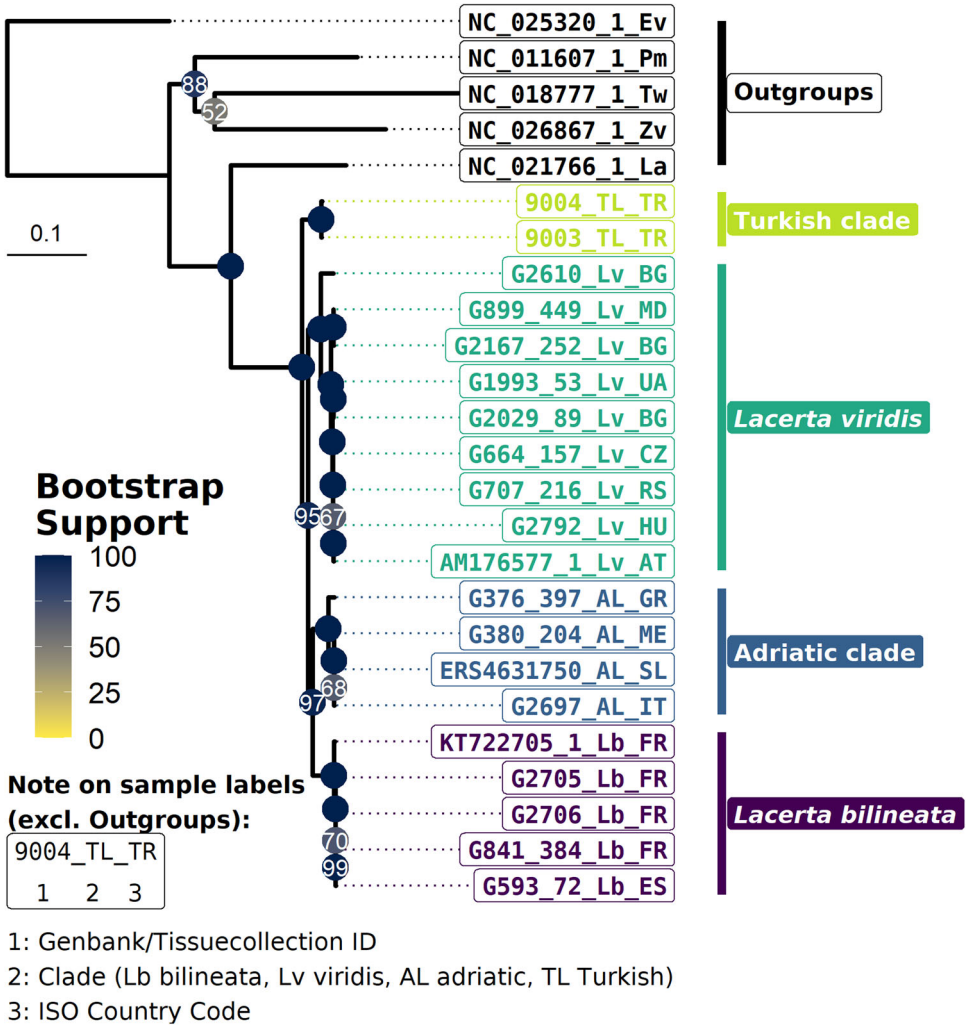

Fig. 3. Maximum Likelihood phylogram of 20 specimens of the *L. viridis* complex based on 13 protein-coding genes plus all rRNA- and tRNA-coding genes. Bootstrap support of phylogram colour-coded as node labels and given on nodes for support values < 100. Outgroup labels refer to the accession numbers in GenBank, specimen labels composed of GenBank/Tissue collection IDs, lineage and geographic location.

highly supported with 97% bootstrap support. Constrained topological analyses revealed significantly depleted likelihood scores and therefore rejected the enforced monophyly of *L. viridis* + Turkish clade (p.AU = 0.0314) and *L. viridis* + Adriatic clade (p.AU = 0.000503), respectively (Fig. S4).

Within *L. bilineata*, two samples from the same collection site (Mâlain, France) were found to be quite different; one of the two mitogenomes was determined in the present study (G2706_Lv_FR) and the other in a previous publication (KT722705.1, Kolora et al., 2017). Resequencing of the previously processed sample revealed differences with respect to the ATP6 gene (KT722705.2). There were idiosyncrasies in the published mitogenome that did not occur in any other *L. bilineata* mitogenome of the present study and a Crimean specimen (Kehlmaier et al., 2020). On closer inspection, the differences in KT722705.1 arose from low coverage regions in the reference-guided mitogenome assembly, which could be due to mapping bias in this region. The resequenced mitogenome (G2705_Lb_FR, KT722705.2) grouped close to the sample from the same collection site (Fig. 4, supplementary figures S1, S2). Another unexpected finding was that the Italian sample (G2697_Lb_IT) clustered with maximum support within the Adriatic clade. According to its collection site, this sample was within the distribution area of *L. bilineata*.

Although bootstrap and supporting values were high, the pairwise distances were low between the terminal branches. In addition, a pairwise comparison of genetic distances of our samples with more distantly related

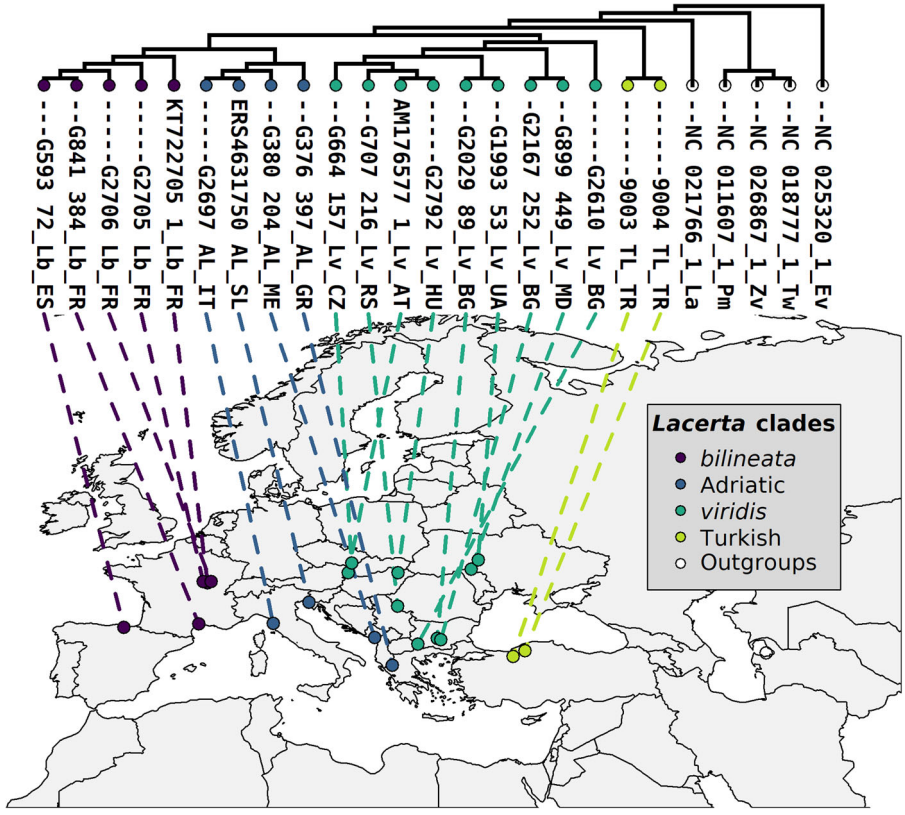

Fig. 4. Maximum Likelihood tree of 20 specimens of the *L. viridis* complex projected on a geographical map. Points represent sampling sites, which are further indicated by the last two characters (country codes) on the sample labels.

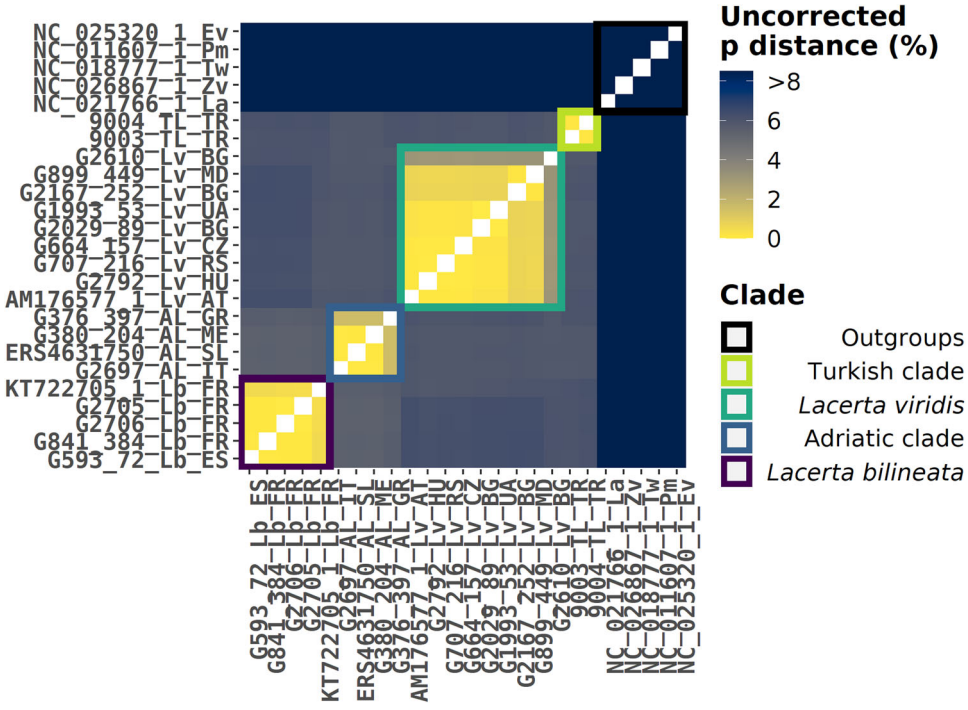

Fig. 5. Heatmap of uncorrected pairwise distances based on the nucleotide alignment of all 13 protein-coding genes for 20 specimens of the *L. viridis* complex. Yellow represents little or no pairwise distance, grey-brown colours represent medium distances and dark blue high distances. Intra-clade distances highlighted by coloured squares.

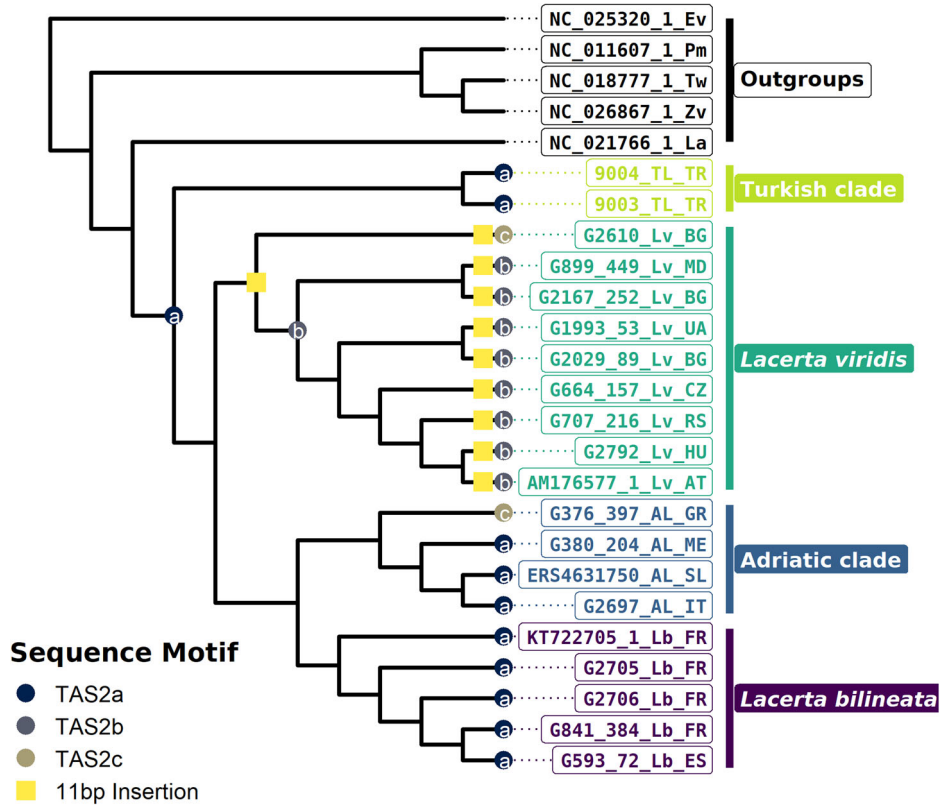

Fig. 6. Hypothetical sequence motif evolution in the *L. viridis* complex mapped to the Maximum Likelihood cladogram. Sequence motifs include three variants of TAS2 as well as an 11 bp insertion within the control region.

lacertids (*L. agilis, Zootoca vivipara, Podarcis muralis, Takydromus wolteri, Eremias vermiculata*) revealed markedly lower genetic distances between the four clades of the *L. viridis* complex than the outgroup representatives compared with one another or with the clades of the *L. viridis* complex (∼2%-∼6% and >10% respectively, Fig. 5).

Discussion

Clade-specific signatures in the control region

All individuals of the *Lacerta viridis* clade within the *L. viridis* complex possess an 11 bp insertion downstream of CSB-1 that was missing in all other samples. This insertion was also found in another study exclusively in *L. viridis*, but not in *L. bilineata*, the Adriatic clade or the outgroup representatives *L. agilis* and *Teira dugesii* (Böhme et al., 2006). In our study, this insertion was also absent in *L. agilis, Zootoca vivipara, Podarcis muralis* and *Takydromus wolteri* (tribe Lacertini), but present in the more distantly related *Eremias vermiculata* (tribe Eremiadini) (systematics follows Arnold et al., 2007). Thus, with some caution, this insertion might be a specific diagnostic molecular signature for the *L. viridis* clade within the *L. viridis* complex. In addition, three variants of the Termination Associated Sequence TAS 2 were found within the *L. viridis* complex (Table 1). TAS-2a was present in the *L. bilineata* and the Adriatic and Turkish clades, whereas TAS-2b was characteristic for the *L. viridis* clade. However, there were two prominent exceptions that have the third variant TAS-2c: a Bulgarian sample of *L. viridis* (G2610_Lv_BG), which is sister to all other samples in the *L. viridis* clade, and a Greek sample of the Adriatic clade (G376_AL_GR). An evolutionary scenario of these (mostly) clade-specific sequence motifs in the CR is shown in Fig. 6 based on maximum parsimony by hand. The distribution of TAS-2c within the *L. viridis* complex limits the diagnostic value of TAS. Therefore, and in face of the limited body of evidence, we do not want to draw extensive conclusions on the taxonomic values of the signatures detected here and in studies before. Also, for another gene, diagnostic signatures have been previously suggested. Godinho et al. (2005) stated that intron 7 of the nuclear β-fibrinogen gene is diagnostic for discriminating *L. bilineata* and *L. viridis*. However, studying more samples, this locus turned out to be highly polymorphic and insufficient for species-

diagnostic purposes (Marzahn et al., 2016). Nevertheless, it is worth to track the detected TAS sequence motifs and search for other sequence motifs, because they might turn out to be taxonomically useful for disentangling the different clades and shed additional light on the phylogeny within this species complex.

Phylogenetic analyses

Our phylogenetic analyses of 20 mitogenomes resolved the branching pattern of the four divergent clades of the *L. viridis* complex for the first time. The recently discovered Turkish clade (Marzahn et al., 2016), sister to all other members of the complex, has to be nomenclaturally identified with the subspecies *L. viridis meridionalis* Cyrén, 1933 (Marzahn et al., 2016), while many populations previously assigned to this subspecies from the southern Balkans represent *L. viridis viridis* sensu lato.

A longstanding issue has been the phylogenetic relationship of the Adriatic clade, whether it is more closely related to *L. bilineata* or to *L. viridis* (Böhme et al., 2007b; Joger et al., 2007; Marzahn et al., 2016). Our data provide firm evidence that the mitochondrial lineage of this clade is closer related to *L. bilineata*.

An unexpected and striking finding was that one sample from Calambrone, Pisa province (Italy), which is within the distribution area of *L. bilineata*, clustered with the Adriatic clade. In order to rule out any confusion, we sequenced the cytb gene of another sample from the same collection site (ERS4631750). In exploratory analyses, this second sample also clustered with the Adriatic clade. This finding calls for further research. The easternmost records of *L. bilineata* are known from the island of Cres, Croatia (Böhme et al., 2007b; Brückner et al., 2002; Marzahn et al., 2016). However, the occurrence of *L. bilineata* on Cres seems to be isolated and encircled by the distribution range of the Adriatic clade, so that introduction has been considered for the green lizards of Cres (Speybroeck et al., 2020). Also, the former occurrence of *L. bilineata* on the southern Crimean Peninsula has been recently shown to result from introduction (Kehlmaier et al., 2020). The westernmost records for the Adriatic clade, on the other hand, are from Friuli-Venezia Giulia, northeastern Italy (Marzahn et al., 2016), approximately 330 km from the record in Pisa province (Tuscany) and beyond the Apennine Mountains. Given that sample confusion can be excluded, this suggests that either all previous studies overlooked a massive westernmost range extension of the Adriatic clade or that the occurrence of the Adriatic clade in Pisa province refers to translocated animals.

Are the four divergent clades distinct species?

Based on hybridization experiments, Rykena (1991) suggested species status for western *L. bilineata* (Daudin, 1802) and eastern *L. viridis* (Laurenti, 1768) due to a decreased hatching success across several captive-bred generations. Even though this conclusion has been controversial for three decades (Arnold & Owenden, 2002; Böhme et al., 2007b; Brückner et al., 2002; Godinho et al., 2005; Mayer & Beyerlein, 2001), both taxa have been largely recognized as distinct species (Speybroeck et al., 2020). According to allozyme studies, gene flow was assumed to occur predominantly unidirectionally from *L. viridis* into *L. bilineata* in the putative contact zone in northeastern Italy and Slovenia (Amann et al., 1997; Joger et al., 2001). However, later studies (Böhme et al., 2007b; Marzahn et al., 2016) showed this region to be occupied by the Adriatic lineage and not by *L. bilineata* as previously assumed. Based on whole genome and transcriptome studies of two green lizards, a *L. bilineata* from Mâlain, France (G2705_Lv_FR), and from Tokay, Hungary G2792_Lv_HU), suggested that gene flow was primarily unidirectional from *L. bilineata* to *L. viridis* after their split at least 1.15 million years ago (Kolora et al., 2019), i.e. in the opposite direction as suggested by the allozyme studies.

A pairwise comparison of genetic distances of our samples with other lacertids revealed markedly lower genetic distances between the four clades than compared to the outgroup representatives and the distances among the latter (Fig. 5). However, a comparison of genetic distances based on 12 S rRNA and 16 S rRNA genes among samples of the *L. viridis* and the *L. bilineata* clades and among samples of two other closely related *Lacerta* species, *L. trilineata* and *L. pamphylica,* show that they are within the same range (Mayer & Beyerlein, 2001). Although genetic distances, in particular of mtDNA, per se cannot be used to delimit species, our findings do not unambiguously support species status for the four clades within the *L. viridis* complex and call for further research that tackles nuclear genomic variation of all four lineages and the viability of hybrids and gene flow within putative hybrid zones.

Conclusion

This study resolves for the first time the phylogeny of the four mitochondrial clades of the *L. viridis* complex. To some extent, clade-specific sequence motifs indicate divergent evolution in the control region among these clades. The genetic distances between the mitogenomes of the four clades were markedly lower than the

divergences between the four clades and the congeneric species *L. agilis*, casting doubts on the species status of *L. bilineata*, *L. viridis* and the two unnamed clades (Adriatic clade, Turkish clade). Nevertheless, the genetic divergences suggest that the four clades are taxonomically distinct and could represent distinct subspecies. A record of the Adriatic clade in northwestern Italy (Pisa province), within the distribution range of *L. bilineata*, calls for further research.

Acknowledgements

We thank the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig for financial support.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Supplemental data

Supplemental data for this article can be accessed here: https://doi.org/10.1080/14772000.2021.1912205.

Data accessibility statement

Mitochondrial genomes have been submitted to the European Nucleotide Archive under the Bioproject number PRJEB40158, sample number ERS5043509. Code, plots and bioinformatic pipeline used in this study are available at https://github.com/RJauss/LacertaViridisMitogenomes.

ORCID

Uwe Fritz http://orcid.org/0000-0002-6740-7214

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Altekar, G., Dwarkadas, S., Huelsenbeck, J. P., & Ronquist, F. (2004). Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics (Oxford, England)*, *20*, 407–415. https://doi.org/10.1093/bioinformatics/btg427

Amann, T., Rykena, S., Joger, U., Nettmann, H. K., & Veith, M. (1997). Zur artlichen Trennung von *Lacerta bilineata* Daudin, 1802 und *L. viridis* (Laurenti, 1768). *Salamandra*, *33*, 255–268.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arnold, E. N., Arribas, O., & Carranza, S. (2007). Systematics of the Palaearctic and Oriental lizard tribe Lacertini (Squamata: Lacertidae: Lacertinae), with descriptions of eight new genera. *Zootaxa*, *1430*, 1–86. https://doi.org/10.11646/zootaxa.1430.1.1

Arnold, E. N., & Owenden, D. W. (2002). *Reptiles and amphibians of Europe (2nd ed.). Princeton field guides*. Princeton University Press.

Bernt, M., Bleidorn, C., Braband, A., Dambach, J., Donath, A., Fritzsch, G., Golombek, A., Hadrys, H., Jühling, F., Meusemann, K., Middendorf, M., Misof, B., Perseke, M., Podsiadlowski, L., Reumont, B. v., Schierwater, B., Schlegel, M., Schrödl, M., Simon, S., … Struck, T. H. (2013a). A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny. *Molecular Phylogenetics and Evolution*, *69*, 352–364. https://doi.org/10.1016/j.ympev.2013.05.002

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., & Stadler, P. F. (2013b). Mitos: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*, 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Bernt, M., Merkle, D., Ramsch, K., Fritzsch, G., Perseke, M., Bernhard, D., Schlegel, M., Stadler, P. F., & Middendorf, M. (2007). Crex: Inferring genomic rearrangements based on common intervals. *Bioinformatics (Oxford, England)*, *23*, 2957–2958. https://doi.org/10.1093/bioinformatics/btm468

Böhme, M. U., Fritz, U., Kotenko, T., Džukić, G., Ljubisavljević, J., Tzankov, N., & Berendonk, T. U. (2007b). Phylogeography and cryptic variation within the *Lacerta viridis* complex (Lacertidae, Reptilia). *Zoologica Scripta*, *36*, 119–131. https://doi.org/10.1111/j.1463-6409.2006.00262.x

Böhme, M. U., Fritzsch, G., Tippmann, A., Schlegel, M., & Berendonk, T. U. (2007a). The complete mitochondrial genome of the green lizard *Lacerta viridis viridis* (Reptilia: Lacertidae) and its phylogenetic position within squamate reptiles. *Gene*, *394*, 69–77. https://doi.org/10.1016/j.gene.2007.02.006

Böhme, M. U., Schneeweiss, N., Fritz, U., Moravec, J., Majlath, I., Majlatova, V., & Berendonk, T. U. (2006). Genetic differentiation and diversity of *Lacerta viridis viridis* (Laurenti, 1768) within the northern part of its range: an investigation using mitochondrial haplotypes. *Salamandra*, *42*, 29–40.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brückner, M., Klein, B., Düring, A., Mentel, T., Rabus, S., & Soller, J. T. (2002). Phylogeographische Analyse des *Lacerta viridis/bilineata* Komplexes: Molekulare Muster und Verbreitung. *Mertensiella*, *13*, 45–51.

Dabney, J., & Meyer, M. (2012). Length and GC-biases during sequencing library amplification: A comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechniques*, *52*, 87–94. https://doi.org/10.2144/000113809

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). Prottest 3: Fast selection of best-fit models of protein

evolution. *Bioinformatics*, *27*, 1164–1165. https://doi.org/10.1093/bioinformatics/btr088

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*, 772. https://doi.org/10.1038/nmeth.2109

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, *39*, 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Fritzsch, G., Schlegel, M., & Stadler, P. F. (2006). Alignments of mitochondrial genome arrangements: Applications to metazoan phylogeny. *Journal of Theoretical Biology*, *4*, 511–520.

Godinho, R., Crespo, E., Ferrand, N., & Harris, D. J. (2005). Phylogeny and evolution of the green lizards, *Lacerta* spp. (Squamata: Lacertidae) based on mitochondrial and nuclear DNA sequences. *Amphibia-Reptilia*, *26*, 271–285. https://doi.org/10.1163/156853805774408667

Goto, N., Prins, P., Nakao, M., Bonnal, R., Aerts, J., & Katayama, T. (2010). Bioruby: Bioinformatics software for the Ruby programming language. *Bioinformatics (Oxford, England)*, *26*, 2617–2619. https://doi.org/10.1093/bioinformatics/btq475

Gustinich, S., Manfioletti, G., Del Sal, G., Schneider, C., & Carninci, P. (1991). A fast method for high quality genomic DNA extraction from whole human blood. *BioTechniques*, *11*, 298–302.

Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads a baiting and iterative mapping approach. *Nucleic Acids Research*, *41*, e129–e129. https://doi.org/10.1093/nar/gkt371

Hahn, C., Schultz, D. T. (2018). Chrishah/Mitobim: Version 1.9.1. Zenodo. https://zenodo.org/record/1408102#.Xnouf3K1I2w

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907–913. https://doi.org/10.1038/35016000

Joger, U., Amann, T., & Veith, M. (2001). Phylogeographie und genetische Differenzierung im *Lacerta viridis/bilineata* Komplex. *Mertensiella*, *13*, 60–68.

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.

Katoh, K., & Standley, D. M. (2013). Mafft multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kehlmaier, C., Zinenko, O., & Fritz, U. (2020). The enigmatic Crimean green lizard (*Lacerta viridis magnifica*) is extinct but not valid: Mitogenomics of a 120-year-old museum specimen reveals historical introduction. *Journal of Zoological Systematics and Evolutionary Research*, *58*, 303–307. https://doi.org/10.1111/jzs.12345

Kolora, S. R., Faria, R., Weigert, A., Schaffer, S., Grimm, A., Henle, K., Sahyoun, A. H., Stadler, P. F., Nowick, K., Bleidorn, C., & Schlegel, M. (2017). The complete mitochondrial genome of *Lacerta bilineata* and comparison with its closely related congener *L. viridis*. *Mitochondrial DNA Part A*, *28*, 116–118. https://doi.org/10.3109/19401736.2015.1111349

Kolora, S. R. R., Weigert, A., Saffari, A., Kehr, S., Walter Costa, M. B., Spröer, C., Indrischek, H., Chintalapati, M., Lohse, K., Doose, G., Overmann, J., Bunk, B., Bleidorn, C., Grimm-Seyfarth, A., Henle, K., Nowick, K., Faria, R., Stadler, P. F., & Schlegel, M. (2019). Divergent evolution in the genomes of closely related lacertids, *Lacerta viridis* and *L. bilineata*, and implications for speciation. *GigaScience.*, *8*, 1–15.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*, 1547–1549. https://doi.org/10.1093/molbev/msy096

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359. https://doi.org/10.1038/nmeth.1923

Liu, P., Zhu, D., Zhao, W.-G., & Ji, X. (2016). The complete mitochondrial genome of the common lizard *Zootoca vivipara* (Squamata: Lacertidae). *Mitochondrial DNA, Part A*, *27*, 1944–1945.

Marzahn, E., Mayer, W., Joger, U., Ilgaz, Ç., Jablonski, D., Kindler, C., Kumlutaş, Y., Nistri, A., Schneeweiss, N., Vamberger, M., Žagar, A., & Fritz, U. (2016). Phylogeography of the *Lacerta viridis* complex: mitochondrial and nuclear markers provide taxonomic insights. *Journal of Zoological Systematics and Evolutionary Research*, *54*, 85–105. https://doi.org/10.1111/jzs.12115

Mayer, W., & Beyerlein, P. (2001). Genetische Differenzierung des *Lacerta viridis/bilineata* Komplexes und von *Lacerta trilineata* in Griechenland: Mitochondriale DNA-Sequenzen. *Mertensiella*, *31*, 52–59.

Mendes, J., Harris, D. J., Carranza, S., & Salvi, D. (2016). Evaluating the phylogenetic signal limit from mitogenomes, slow evolving nuclear genes, and the concatenation approach. New insights into the Lacertini radiation using fast evolving nuclear genes and species trees. *Molecular Phylogenetics and Evolution*, *100*, 254–267. https://doi.org/10.1016/j.ympev.2016.04.016

Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., & Marshall, D. (2010). Tablet – next generation sequence assembly visualization. *Bioinformatics (Oxford, England)*, *26*, 401–402. https://doi.org/10.1093/bioinformatics/btp666

Nettmann, H.-K. (2001). Die Smaragdeidechsen (*Lacerta* s. str.) – Eine Übersicht über Verwandtschaft und Formenvielfalt. *Mertensiella*, *13*, 11–32.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*, 268–274. https://doi.org/10.1093/molbev/msu300

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Perseke, M., Fritzsch, G., Ramsch, K., Bernt, M., Merkle, D., Middendorf, M., Bernhard, D., Stadler, P. F., & Schlegel, M. (2008). Evolution of mitochondrial gene orders in echinoderms. *Molecular Phylogenetics and Evolution*, *47*, 855–864. https://doi.org/10.1016/j.ympev.2007.11.034

Perseke, M., Golombek, A., Schlegel, M., & Struck, T. H. (2013). The impact of mitochondrial genome analyses on the understanding of deuterostome phylogeny. *Molecular Phylogenetics and Evolution*, *66*, 898–905. https://doi.org/10.1016/j.ympev.2012.11.019

Podnar, M., Haring, E., Pinsker, W., & Mayer, W. (2007). Unusual origin of a nuclear pseudogene in the Italian wall lizard: Intergenomic and interspecific transfer of a large section of the mitochondrial genome in the genus *Podarcis* (Lacertidae). *Journal of Molecular Evolution*, *64*, 308–320. https://doi.org/10.1007/s00239-005-0259-0

Qin, P.-S., Tao, C.-R., Yin, S., Li, H.-M., Zeng, D.-L., & Qin, X.-M. (2014). Complete mitochondrial genome of *Lacerta agilis* (Squamata, Lacertidae). *Mitochondrial DNA*, *25*, 416–417. https://doi.org/10.3109/19401736.2013.809436

R Core Team. (2019). R: A Language and Environment for Statistical Computing. https://www.R-project.org/

Revell Liam, J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

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. *Systematic Biology*, *61*, 539–542. https://doi.org/10.1093/sysbio/sys029

Rykena, S. (1991). Kreuzungsexperimente zur Prüfung der Artgrenzen im Genus *Lacerta* sensu stricto. *Mitteilungen aus dem Museum für Naturkunde in Berlin*, *67*, 55–68. https://doi.org/10.1002/mmnz.19910670108

Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. *Systematic Biology*, *51*, 492–508. https://doi.org/10.1080/10635150290069913

Speybroeck, J., Beukema, W., Dufresnes, C., Fritz, U., Jablonski, D., Lymberakis, P., Martínez-Solano, I., Razzetti, E., Vamberger, M., Vences, M., Vörös, J., & Crochet, P.-A. (2020). Species list of the European herpetofauna – 2020 update by the Taxonomic Committee of the Societas Europea Herpetologica. *Amphibia-Reptilia*, *41*, 139–189. https://doi.org/10.1163/15685381-bja10010

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics (Oxford, England)*, *22*, 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *4*, 453–464.

Tong, Q.-L., Yao, Y.-T., Lin, L.-H., & Ji, X. (2016). The complete mitochondrial genome of *Eremias vermiculata* (Squamata: Lacertidae). *Mitochondrial DNA, Part A*, *27*, 1447–1448. https://doi.org/10.3109/19401736.2014.953086

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Yu, D.-N., & Ji, X. (2013). The complete mitochondrial genome of *Takydromus wolteri* (Squamata: Lacertidae). *Mitochondrial DNA*, *24*, 3–5. https://doi.org/10.3109/19401736.2012.710223

Associate Editor: David Gower