Available online at www.sciencedirect.com
ScienceDirect
Gene 394 (2007) 69–77
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

The complete mitochondrial genome of the Green Lizard *Lacerta viridis viridis* (Reptilia: Lacertidae) and its phylogenetic position within squamate reptiles

M.U. Böhme [a,*], G. Fritzsch [b], A. Tippmann [a], M. Schlegel [a], T.U. Berendonk [a]

[a] *University of Leipzig, Biology II, Molecular Evolution and Animal Systematics, Talstrasse 33, 04103 Leipzig, Germany*
[b] *Interdisciplinary Centre for Bioinformatics (IZBI), University of Leipzig, Härtelstrasse 16-18, 04107 Leipzig, Germany*

Received 20 December 2006; accepted 9 February 2007
Available online 23 February 2007

Abstract

For the first time the complete mitochondrial genome was sequenced for a member of Lacertidae. *Lacerta viridis viridis* was sequenced in order to compare the phylogenetic relationships of this family to other reptilian lineages. Using the long-polymerase chain reaction (long PCR) we characterized a mitochondrial genome, 17,156 bp long showing a typical vertebrate pattern with 13 protein coding genes, 22 transfer RNAs (tRNA), two ribosomal RNAs (rRNA) and one major noncoding region. The noncoding region of *L. v. viridis* was characterized by a conspicuous 35 bp tandem repeat at its 5′ terminus. A phylogenetic study including all currently available squamate mitochondrial sequences demonstrates the position of Lacertidae within a monophyletic squamate group. We obtained a narrow relationship of Lacertidae to Scincidae, Iguanidae, Varanidae, Anguidae, and Cordylidae. Although, the internal relationships within this group yielded only a weak resolution and low bootstrap support, the revealed relationships were more congruent with morphological studies than with recent molecular analyses.

Keywords: Reptilia; Squamata; Lacertidae; Mitochondrial genome; Control region; Minisatellite; Repeat unit

1. Introduction

Vertebrate mitochondrial genomes are circular molecules with a characteristic length of 16–19 kb. Generally, the mt genome comprises 13 protein coding genes, 22 tRNAs, two rRNAs and one major noncoding region, the control region (reviewed e.g. in Boore, 1999; Pereira, 2000). The replication of this double stranded molecule takes place asymmetrically for the heavy and light strand, starting from two independent replication origins (O_H, O_L). This fact and the impact of a highly reactive environment within the organelle are supposed to cause high mutation rates, which may conceivably differ between both strands within the mitochondrial genome (Reyes et al., 1998; Rand, 2001).

Until today, 541 complete vertebrate mitochondrial genomes are available in GenBank (status 03/2006). Within this mass of information only 24 complete squamate genomes exist (Janke et al., 2001; Kumazawa, 2004; Amer and Kumazawa, 2005; Macey et al., 2005). Most of them belong to members of Amphisbaenia (Macey et al., 2004) and Serpentes (Kumazawa et al., 1998; Kumazawa, 2004; Dong and Kumazawa, 2005). From the large infraorder of Scincomorpha, including Lacertoidea, Scincoidea and Teiioidea, only two mitochondrial genomes have been submitted to GenBank so far (Kumazawa and Nishida, 1999; Kumazawa, 2004). In this study we characterize the structural features of one of the first mitochondrial genomes

Abbreviations: ATP6 and ATP8, genes for ATP synthase subunits 6 and 8; bp, base pair; COI–III, genes for cytochrome *c* oxidase subunits I–III; Cyt *b*, gene for cytochrome *b*; CR, major noncoding region, control region; CSB, conserved sequence blocks of control region; kb, kilo base pairs; ML, Maximum Likelihood; MP, Maximum Parsimony; mtDNA, mitochondrial DNA; mt genome, mitochondrial genome; ND1–6, genes for NADH dehydrogenase subunits 1–6; NJ, Neighbor Joining; O_L and O_H, origin of replication light strand and heavy strand of mitochondrial DNA; PCR, polymerase chain reaction; R1 and R2 repeat sequences within control region; 12S rRNA and 16S rRNA, genes for small and large subunit of ribosomal RNA; TAS, termination associated sequences; tRNAX, transfer RNA genes with corresponding amino acids denoted by a one letter code.; tRNAxxx, transfer RNA genes with corresponding amino acids denoted by a three letter code.

* Corresponding author. Tel.: +49 3419736743; fax: +49 3419736789.
E-mail address: maboehme@rz.uni-leipzig.de (M.U. Böhme).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.02.006

of a representative of the family Lacertidae (Lacertoidea). The Green Lizard *Lacerta viridis* (Lacertidae, *Lacerta*) is distributed over Middle and Eastern Europe and is currently subdivided into five subspecies. Because of this complex situation we chose a member of the nominate form *L. v. viridis* for characterizing the mt genome of this species.

In reference to the still controversial data about higher level relationships within the order Squamata, the group of Lacertidae is of particular interest. Therefore, we analyzed the phylogenetic position of the Lacertidae within the squamate lineages by using the full information of the protein coding mitochondrial genes, consisting of 13 genes and all codon positions. For this, we used all currently available 25 mitochondrial genomes of squamate reptiles, including *Varanus komoensis* (Kumazawa and Endo, 2004).

2. Materials and methods

2.1. Lizard samples and amplification of mitochondrial DNA

Extracted mitochondrial DNA of *L. v. viridis* from Hundsheim (Austria) was kindly provided by Werner Mayer (Naturhistorische Sammlung, Wien). The amplification of the complete mitochondrial genome followed several steps. Primer sequences are available in the supplemental material. First, primers LvF and 12L (LvF, Böhme et al., 2006; 12L, Brehm et al., 2003) were used to obtain a fragment, which encompassed the main part of cytochrome *b*, tRNA Thr, tRNA Pro, complete control region and part of 12S rRNA (Fig. 1). Using the primers LvF3 and H3056 (H3056, Hedges, 1994) we were able to amplify parts of the 12S rRNA and 16S rRNA genes (Fig. 1). These fragments were sequenced and the primer LvR4 was designed. The primers ND4R, L2602 and ND4F were modified according to the recently published primers and sequences of other reptiles (Hedges, 1994; Arevalo et al., 1994; Busack et al., 2005). The primer pair L2602 and ND4R was used to perform a long PCR, which compassed a 8800 bp fragment from the 16S rRNA to the ND4 gene (Fig. 1). Also employing a long PCR a 3750 bp long mitochondrial fragment spanning from the ND4 gene to the cytochrome *b* gene was amplified, using the primer pair ND4F and LvR4 (Fig. 1). PCR amplifications were accomplished under following conditions: 1× reaction buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 15 pmol of each primer and 1 U Taq polymerase (Sigma). The PCR was performed on an Eppendorf Mastercycler starting with 95 °C for 2 min and 35 cycles at 94 °C for 30 s, 50 °C or 55 °C (12S–16S) for 1 min, 72 °C for 3 min. For the long PCR amplifications we used the Expand High Fidelity System (Roche) under the following conditions: 1× reaction buffer 1, 1.5 mM $MgCl_2$, 0.2 mM dNTPs,

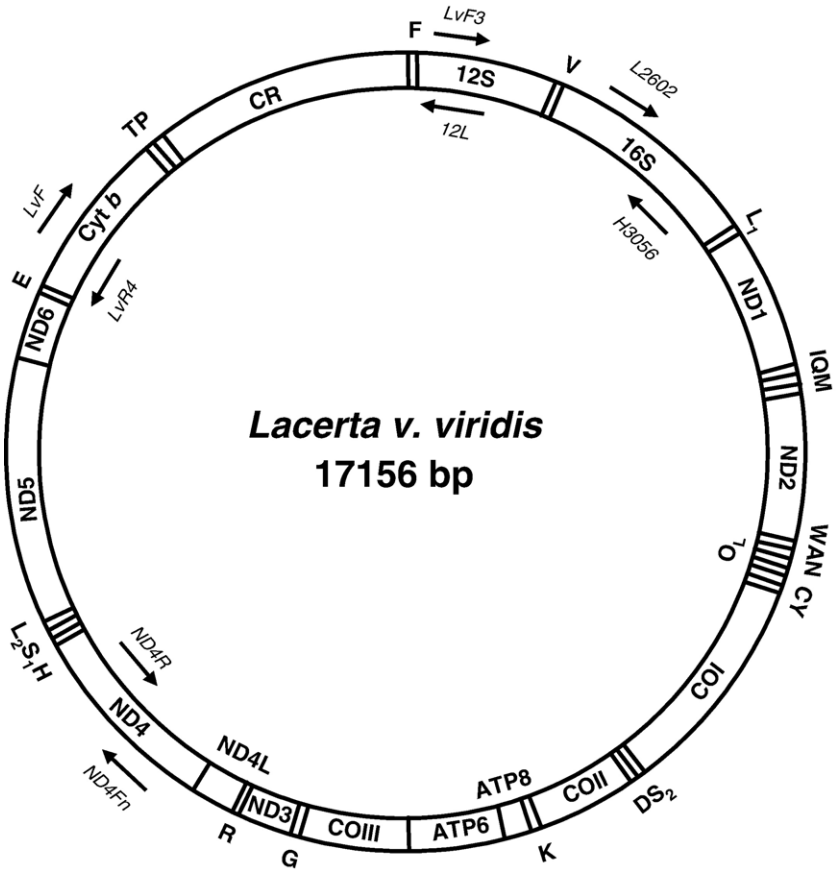

Fig. 1. The organization of the *L. v. viridis* mitochondrial genome. tRNAs are denoted with single letter amino acid code. The locations of the eight amplification primers are shown with small arrows. Protein coding genes are indicated as follows: ND1–6, NADH dehydrogenase subunits 1–6; COI–III, cytochrome *c* oxidase subunits I–III; ATP6 and ATP8, ATPase subunits 6 and 8; Cyt *b*, cytochrome *b*; CR — major noncoding region, control region.

20 pmol of each primer and 5 units of polymerase (Roche). PCR was performed on an Eppendorf Mastercycler starting with 94 °C for 2 min, 10 cycles at 94 °C for 15 s, 46 °C for 30 s, 68 °C for 8 min and 10 cycles at 94 °C for 15 s, 46 °C for 30 s, 68 °C for 8 min 50 s. A final step followed with 68 °C for 7 min and 4 °C for 10 min. PCR products were checked on 1% agarose gel and purified using Rapid PCR Purification System (Marligen). Internal sequences were obtained by designing nested sequencing primers for "primer walking" within amplified products. These primers will be available upon request and were designed using the Fast PCR v. 3.5.52 (Kalendar, 2005). PCR fragments were sequenced by cycle sequencing, using Terminator Ready Reaction Mix 'Big Dye' v. 3.1 (Applied Biosystems) following the manufacturer's protocol and analyzed using an ABI 3100 DNA Sequencer.

2.2. Sequence analysis

All sequences were checked and corrected by visual inspection using the program Bioedit v. 7.0.1 (Hall, 1999) and aligned using DNASIS v. 7.0 (Hitachi Software Engineering). The locations of the 13 protein coding genes, tRNAs and rRNAs were determined by comparisons with other reptile sequences available in GenBank. Start and stop codon positions of the coding genes were identified using the vertebrate mitochondrial code and amino acid reading frames were checked. Boundaries of noncoding regions were recognized by the range of the coding genes. tRNAs were folded using tRNA Scan v. 1.1 (Lowe and Eddy, 1997) and the Vienna RNA Package (Hofacker, 2003).

2.3. Phylogenetic analysis

To determine the phylogenetic position of *L. v. viridis* sequences of 13 protein coding mitochondrial genes were aligned. In contrast to recent molecular studies (e.g. Kumazawa, 2004, 2007; Dong and Kumazawa, 2005; Douglas et al., 2006; Zhou et al., 2006), we included the light strand coded gene ND6, as we found that neither the ex- or inclusion had an impact on the resulting phylogeny. Complete nucleotide sequences and amino acid sequences of *L. v. viridis* were aligned to additional 31 available complete mt genome sequences (including outgroup members) from GenBank: *Pogona vitticeps* (NC_006922), *Iguana iguana* (NC_002793), *Sceloporus occidentalis* (NC_005960), *Bipes canaliculatus* (NC_006288), *Bipes biporus* (NC_006287), *Bipes tridactylus* (NC_006286), *Geocalamus acutus* (NC_006285), *Amphisbaena schmidti* (NC_006284), *Rhineura floridana* (NC_006282), *Diplometopon zarudnyi* (NC_006283), *Abronia graminea* (NC_005958), *Shinisaurus crocodilurus* (NC_005959), *Teratoscincus keyserlingii* (NC_007008), *Gekko gecko* (NC_007627), *Cordylus warreni* (NC_005962), *Eumeces egregius* (NC_000888), *Varanus komodoensis* (AB_080276, AB_080275), *Acrochordus granulatus* (NC_007400), *Ovophis okinavensis* (NC_007397), *Dinodon semicarinatus* (NC_001945), *Xenopeltis unicolor* (NC_007402), *Cylindrophis ruffus* (NC_007401), *Python regius* (NC_007399), *Boa constrictor* (NC_007398), *Leptotyphlops dulcis* (NC_005961), *Alligator mississippiensis* (NC_001922), *Caiman crocodilus* (NC_002744), *Falco peregrinus* (NC_000878), *Gallus gallus* (NC_001323), *Chelonia mydas* (NC_000886), *Pelomedusa subrufa* (NC_001947).

All sequences were aligned based on amino acid sequences using Clustal 1.83 implemented in Mega 3.1 (Kumar et al., 2004). The further used nucleotide alignment (11,301 sites) followed this amino acid pattern.

A partition homogeneity test, implemented in PAUP*4.0 b10 (Swofford, 2002) was performed to examine differences between each gene. The result for the combined data set of 13 protein coding genes showed a *P*-value of 0.01, therefore the protein coding data sets were congruent and combining the genes did not reduce the phylogenetic accuracy of the whole data set.

To test for potential saturations and levels of homoplasy within the data set, we employed the g1 statistic (Sokal and Rohlf, 1981) implemented in PAUP*4.0b10 to provide a conventional measure of skewness for tree-length distribution and to determine whether the phylogenetic signal was significantly non-random (Hillis and Huelsenbeck, 1992). The significance of the g1 statistic was assessed using the critical values calculated by Hillis and Huelsenbeck (1992).

The combined data set contained 7843 informative characters and a significant phylogenetic signal above background noise as indicated by the g1 statistic ($g1 = -0.723050$, $P = 0.05$; Hillis and Huelsenbeck, 1992). Our data set passed the PTP and left-skewness (g1) test. Accordingly, 1000 permuted data sets or 100,000 randomly generated trees were used ($P < 0.05$). These tests indicated that the data contain more hierarchical structure than expected for random data (Faith and Cranston, 1991; Hillis and Huelsenbeck, 1992).

DNA substitution rate was calculated with a hierarchical likelihood ratio test using MrModeltest 2.2 (Nylander, 2004). We tested all models for each gene, for all genes the best fit was reported for the GTR model. Consequently, we used the GTR model for the analyses of the entire genome. The selected model was GTR+I+G ($-\ln L = 247788.1875$, Tavaré, 1986) with base frequencies of A=0.3941, C=0.3441, G=0.0651, T=0.1966; proportion of invariable sites of I=0.1915 and a gamma distribution shape parameter of variable sites of G=0.6149. A Neighbor Joining tree (NJ) using the GTR model (Tavaré, 1986) and a Maximum Parsimony tree (MP) with heuristic search using ten stepwise additions of sequences and TBR branch swapping option were also constructed using PAUP*4.0b19 (Swofford, 2002). To test the robustness of NJ and MP bifurcations, bootstrap analyses with 10,000 replicates for NJ and 2000 replicates for MP were performed. Because of the high calculation effort for the Maximum Likelihood (ML) approach within the program PAUP we calculated the ML tree using PHYML v. 2.4.4 (Guindon and Gascuel, 2003). The model of nucleotide substitution revealed by the integrated operation of the program was also GTR+I+G with slightly different values for base frequencies, invariable sites and gamma distribution shape parameter but a higher log likelihood value ($-\ln L = 247596.2380$) than revealed by MrModeltest 2.2. The ML analysis was run with the following parameters: base

Table 1
Comparison of the base contents within the whole mitochondrial genome of *L. v. viridis* to different vertebrate species

Species	Percentage of bases within mtDNA (%)					
	A	C	G	T	A+T	G+C
Lacerta viridis viridis	31.3	27.0	13.3	28.4	59.7	40.3
Dinodon semicarinatus [a]	34.7	27.7	12.2	25.3	60	39.9
Pelomedusa subrufa [b]	34	27	12	27	61	39
Bipes biporus [c]	30	30	14	26	56	44
Iguanidae [d]	33–37	24–33	11–13	21–30	54–67	35–46

[a] Kumazawa et al. (1998).
[b] Zardoya and Meyer (1998).
[c] Macey et al. (1998).
[d] Macey et al. (1997b).

frequencies of A=0.39709, C=0.34344, G=0.06496, T=0.19451; proportion of invariable sites of I=0.123 and a gamma distribution shape parameter of variable sites of G=0.488. The robustness of the ML bifurcations was tested by 200 replications. Furthermore, a Bayesian analysis was performed with MrBayes v. 3.1.2 (Ronquist and Huelsenbeck, 2003). The analysis was run with the best-fit model as inferred by PHYML v. 2.4.4 for 1,000,000 generations, with a sampling frequency of 100 generations. From the 10,000 trees found, we determined a subset of trees for building our consensus tree by inspecting likelihood values of trees saved by MrBayes. The burn-in was set to 5000 trees to ensure that stable likelihood values were achieved.

3. Results and discussion

3.1. Genome organization

The obtained complete sequence of *L. v. viridis* was 17,156 bp long and contained 13 protein coding genes, 22 tRNA genes, two rRNAs genes and 1750 bp of noncoding control region (Fig. 1). The complete sequence will be available in GenBank under the accession number AM176577. The structural arrangement of the protein coding genes, tRNA genes, rRNA genes and control region corresponds to the typical vertebrate pattern reviewed in Boore (1999) and are provided as supplemental material.

Most genes were encoded on the H-strand except ND6 and nine tRNAs. Interestingly, most of the protein coding genes showed partially overlapping sequences but we also found spacing sequences up to 17 bp between the different coding features. The nucleotide composition of the presented mitochondrial genome was 31.3% A, 27% C, 13.3% G and 28.4% of T. Therefore, the overall A+T content is 59.7%, which corresponds well with the values found in other vertebrate species (Janke and Arnason, 1997; Janke et al., 2001; Table 1).

3.2. Protein coding genes

Almost all of the 13 proteins showed the regular start codon ATG except genes CO I and CO II, which started with GTG. Within the stop codons we found a more diverse pattern of codon usage (TAA, AGG, TAG) as well as unfinished stop codons within some genes (ND4, ND4L, COIII, COII, ATP8, ND3), which could be completed by post-transcriptional polyadenylation (Ojala et al., 1980; Ojala et al., 1981). The observed lengths of the protein coding genes and the overall base content of protein coding genes, except ND6 (29.5% T, 28.4% C, 30.5% A, 11.7% G) were similar to other vertebrate genomes. The detected strong bias against G on third codon position is comparable to other studies on vertebrate mt genomes (Janke and Arnason, 1997; Janke et al., 2001).

3.3. Noncoding sequences — control region

The major noncoding region of the *L. v. viridis* mitochondrial genome (Fig. 2), the control region (CR), was located between the tRNAs Proline (tRNA P) and Phenylalanine (tRNA F). The base composition of the CR (30.2% A, 22.7% C, 35.8% T, 11.2% G) is comparable with the average mtDNA base content of *L. v. viridis*. Overall organization of this region is comparable to the situation of other vertebrates. The internal organization of the CR (Fig. 2, Table 2) showed three conserved sequence blocks (CSB 1–3) as reported for several vertebrate species (Walberg and Clayton, 1981; Sbisa et al., 1997) and two perfect repeats R1 and R2 comparable with *Lacerta dugesii* and other reptiles (Brehm et al., 2003). Within the first part (position 350 to position 650) of the CR we identified several fragments

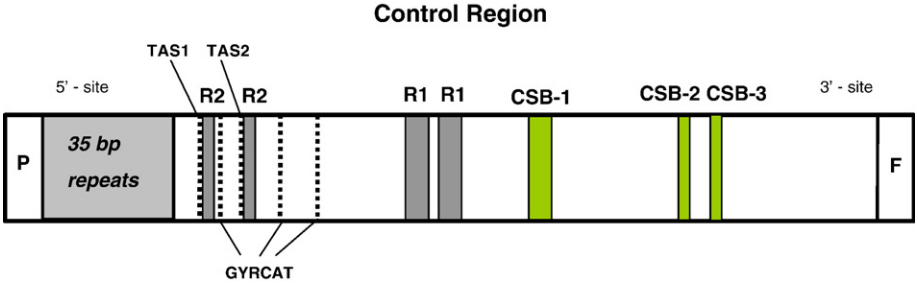

Fig. 2. Feature map of the mitochondrial control region indicating the relative position of the repeat region. Perfect repeats are denoted by R1 and R2. The vertebrate conserved sequence blocks (CSB1–3) and putative TAS sequences and "GYRCAT" motifs (dashed lines) follow the terminology of Brehm et al. (2003). The flanking tRNAs are indicated by a single letter amino acid code.

Table 2
Features of the major noncoding region — control region within the mitochondrial genome of *L. v. viridis*

Feature	Position	Motif (5′–3′)
CSB-1	16438–16475	CTATATGGTATTATTGTCTTAATGCTTGGTAGACATAT
CSB-2	16746–16763	CAAACCCCCCTACCCCCC
CSB-3	16801–16820	TCGCCAAACCCCTAAAACGA
R1 first position	16183–16216	CACCACTTTCTCACTTTTTCCAAGGCCTCTGGTT
R2 first position	15752–15765	ATTTACCCCATGAA
Repeat motif	15405–15439	CACCTGCCGCTTAAAAGCGGCTTTTTTGCCTCCTA
TAS1	15725–15748	ACTATTATGTATATAGTGCATTAA
TAS2	15818–15827	...CATACATTAA...

comparable to termination associated sequences (TAS) or regulative "GYRCAT" motifs within the control region. At the very 5′- beginning of the control region we found several repeats consisting of a 35 bp motif. This minisatellite showed an increase of Cysteine (34%) and a striking decrease of Adenine (17%) in comparison to the rest of the control region, which has also been reported for the minisatellite region in *L. dugesii* (Brehm et al., 2003). Intraspecific studies (Böhme et al., 2006) showed a variable repeat number (6–10 repeats) of this minisatellite causing a variable length of the control region ranging from 1680 bp (6 repeats) up to 1820 bp (10 repeats).

3.4. Ribosomal and transfer RNA genes

Like other mitochondrial genomes the *L. v. viridis* genome contains a small subunit of rRNA (12S) and a large subunit of mt ribosomal rRNA (16S). 12S rRNA (940 bp) was separated by tRNA Valine from the 16S rRNA (1524 bp, Fig. 1).

Within the mitochondrial genome we identified 22 tRNAs ranging from 56 to 79 bp. The arrangement within the mt genome shows no observable differences from a typical vertebrate pattern. All tRNAs were folded and showed typical secondary structures with normal base pairing comparable to other squamate tRNAs. Secondary structures are available as supplemental material. We found the longest spacer of 17 bp between the tRNA Threonine (T) and Proline (P), which could not be assigned to either of them. tRNAs of the so called WANCY cluster displayed the typical organization with an integrated putative origin of replication for the light strand O_L. The assumed O_L was 32 bp long and its sequence motif 5′-GCCTTCTTCCCCCGTTAGAAAAAAAA-CGGGGG-3′ is similar to that of *Eremias grammica* published by Macey et al. (1997a).

3.5. Phylogenetic relationships

To avoid misinterpretations within the following section the nomenclature was adopted from the EMBL Reptile Database (Uetz, 2005). The reconstruction of the phylogenetic relationships is presented by the Bayesian analysis (Fig. 3). Where possible, we added bootstrap values of additional Maximum Likelihood (ML), Maximum Parsimony (MP) and Neighbor Joining (NJ) analyses. The phylogenetic tree in Fig. 3 includes all currently available complete squamate mitochondrial genomes. *Sphenodon punctatus* was not included because of the missing information for the ND5 gene. To allow comparisons with other studies we used members of Crocodylidae, Aves and Testudines as outgroups (e.g. Kumazawa, 2004; Townsend et al., 2004; Dong and Kumazawa, 2005). The widely accepted phylogenetic relationships between these outgroups, in fact that Aves and Crocodylidae form a cluster Archosauria, could not be revealed by any of the used methods. However, a separate phylogenetic analysis consisting only of the outgroup members, showed high support for a Archosauria clade (data not shown). Therefore, we suggest that this unorthodox relationship of Aves and Testudines within our tree is due to a phylogenetic randomization of the outgroups influenced by the ingroup sequences (e.g. Tarrio et al., 2001; Graham et al., 2002). Also within the study of Townsend et al. (2004) the mtDNA data showed such a strange position of Testudines as a sistergroup of Aves and a more narrow relationship of Crocodylidae to the Squamata.

However, we clearly show that the Squamata form a monophyletic group. For the following discussion we have to point out that the phylogenetic pattern especially within the basal splits of Squamata differs between the several analysis methods (Bayes/ML versus MP/NJ). As Bayesian and Maximum Likelihood methods showed the same well supported phylogeny, we assume that this phylogeny was the most likely one and we will discuss this pattern first. Nevertheless, all methods showed a high agreement concerning the relationships within the higher infraorders and superfamilies of the Squamata.

The most important and interesting feature was the position of *L. viridis* within the phylogeny (Fig. 3). We found that Scincidae (**2**) and Lacertidae (**3**) form a monophyletic group, but this relationship had only weak bootstrap support. We were not able to reveal evidence for the taxon of Scincomorpha, as Cordylidae (**7**) appeared as a more distant branch and complete mitochondrial data of further corresponding taxa like Teiioidae and Xantusiidae were not available. What can be concluded from our analyses is that *L. viridis* is a typical member of Lacertidae and its position correlates better with morphological studies (Lee, 2005) than with the recent molecular studies, where Lacertidae were grouped closely to the Amphisbaenidae (Townsend et al., 2004; Vidal and Hedges, 2005; Kumazawa, 2007). This was not the case in any of our analyses.

In agreement with other recent molecular studies (e.g. Townsend et al., 2004; Vidal and Hedges, 2005; Douglas et al., 2006; Kumazawa, 2007; Zhou et al., 2006) we did not find the traditional dichotomy of the basal splits of Iguania and

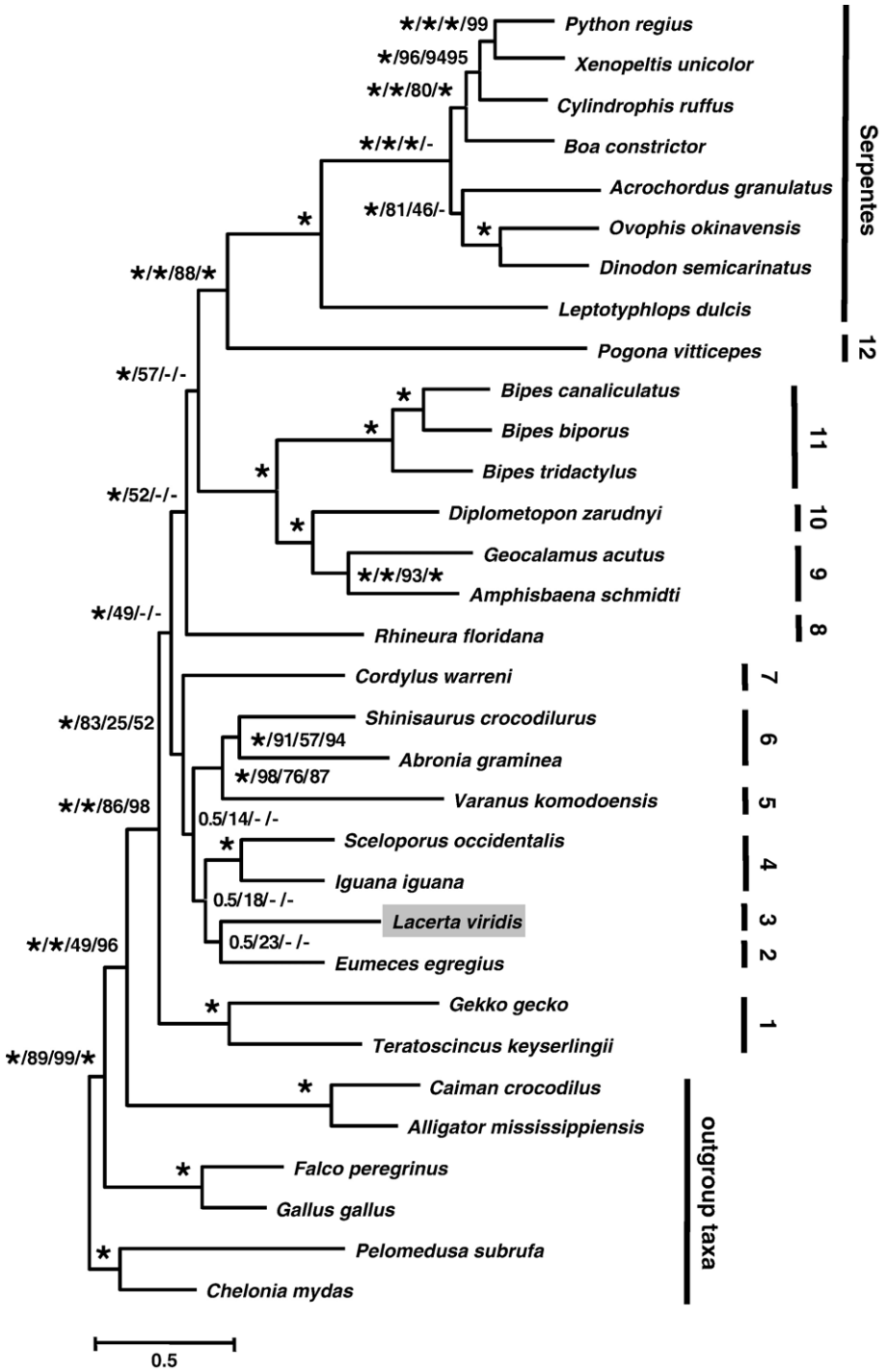

Fig. 3. Bayesian phylogram clarifying the phylogenetic relationships between *Lacerta viridis viridis* as a member of Lacertidae (**3**) and other squamate groups revealed by the analysis of thirteen mitochondrial protein coding genes. Bootstrap values from 200 replications for maximum Likelihood (ML), 2000 replications for Maximum Parsimony (MP) and 10,000 replicates for Neighbor Joining method (NJ) were employed. The first value at the indicated nodes represents the probability value for Bayes, the second the bootstrap value for ML, the third for MP and the fourth for NJ. Asterisks indicate probability values of 1.0 or bootstrap values of 100%. Bifurcations labelled with one asterisk were highly supported by all methods. Dashes indicate that the node was not supported by the specific method. The squamate groups of the different taxonomic levels are indicated by the following numbers: **1** Gekkonidae, **2** Scincidae, **3** Lacertidae, **4** Iguanidae, **5** Varanidae, **6** Anguidae, **7** Cordylidae, **8** Rhineuridae, **9** Amphisbaenidae, **10** Trogonophidae, **11** Bipedidae, **12** Agamidae.

Scleroglossa, in all trees these groups were paraphyletic. In our analysis the monophyly of Iguania was broken by the nested position of monophyletic Iguanidae (**4**) higher within the tree in close relationship to the Lacertidae (**3**) and Scincidae (**2**), and the position of the agamide *P. vitticeps* (**12**) as the sistergroup of the highly supported monophyletic Serpentes (snakes).

The tree presented here shows a well supported basal position of Gekkonidae (**1**) to all other squamate groups, which

is comparable to other molecular studies (Townsend et al., 2004; Vidal and Hedges, 2005; Douglas et al., 2006; Kumazawa, 2007). However, the other main splits within the tree do not agree with most of these studies. Concordant to the study of Douglas et al. (2006) we obtained a major split within the residual squamate groups, which separated the "limb-reduced" taxa of Amphisbaenia (**8–11**) and Serpentes from other lizards. The unusual sistergroup relationship of Serpentes and Agamidae (**12**) was well supported by all analysis methods (ML, NJ, MP, Bayes). This relationship should be interpreted with caution, as it is possible that the position of *P. vitticeps* may be influenced by a long branch effect similar to the mitochondrial based phylogeny in the study of Townsend et al. (2004) and Douglas et al. (2006). Furthermore this position of Agamidae (**12**) could also be affected by the absence of a more related taxon like Chamaeleonidae (Lee, 2005). Recent molecular studies did not reveal such a relationship as they found Agamidae more related to other acrodont groups (Townsend et al., 2004; Kumazawa, 2007). However, the morphological studies of Lee (2005) using osteological data suggested, that a closer relationship of these taxa could be conceivable. The revealed relationships within the serpent taxa were well supported and highly congruent to the results in Dong and Kumazawa (2005). Considering the second "limb-reduced"clade, the Amphisbaenidae (**9**), Trogonophidae (**10**) and Bipedidae (**11**) form a highly supported monophyletic group in agreement with the previous molecular studies of amphisbaenian relationships (Kearney and Stuart, 2004; Macey et al., 2004). The more distant position of Rhineuridae (**8**) basal to all other "limb-reduced" taxa corresponds not well with these studies, which showed the limbless Rhineuridae as the first split of a monophyletic Amphisbaenia group.

Apart from the "limb-reduced" clade, the anguimorph, iguanide and scincomorph representatives appeared as one major group within the tree. The basal relationships within this group were not highly supported. In contrast to the study of Douglas et al. (2006) and Zhou et al. (2006) we found no monophyletic relationship of Scincomorpha as the Cordylidae (**7**) are presented as a single branch distant to the morphologically related Lacertidae (**3**) and Scincidae (**2**). Furthermore, we found a well supported anguimorph clade of Varanidae (**5**) and Anguidae (**6**) which was also revealed by recent molecular studies (e.g. Townsend et al., 2004; Vidal and Hedges, 2004, 2005; Douglas et al., 2006; Zhou et al., 2006). However, this position is in conflict with the morphological studies where Anguimorpha are often associated with snakes (Lee, 2005). Another similarity to recent molecular studies is shown in the nested position of Iguanidae (**4**) within this group, which is highly contrary to morphological data (Lee, 2005). Townsend et al. (2004) suggested that this position could be also due to a long branch effect. But within all of the four trees we found this nested position, therefore we would agree with the conclusions of Vidal and Hedges (2005) that this could also be an evidence for a highly specialized rather than a primitive lineage of the Iguanidae within the Squamata.

Despite the fact that our data set showed no significant levels of homoplasy and saturation, the results demonstrate an instability of the phylogenetic signal if we compare the Bayesian and Maximum Likelihood methods with the Neighbor Joining (NJ) and Maximum Parsimony (MP) methods. MP and NJ phylogenies did not support the basal position of Gekkonidae, revealed by Townsend et al. (2004) and Vidal and Hedges (2005) but supported a basal position of the Serpentes/Agamidae clade to all other squamate groups (data not shown). This scenario would support the suggestion mentioned by Kumazawa (2004) that snakes and lizards evolved independently from each other and Sauria may form a monophyletic group (Zhou et al., 2006). This would be additionally supported by the fact that we also did not find a relationship between Serpentes and Varanidae with molecular methods as it had been suggested on the basis of morphological information (Lee, 2005).

With the MP and NJ trees Gekkonidae emerged later within the tree in a sistergroup relationship to the amphisbaenian taxa, which is a conceivable relationship according to the analyses of soft anatomies (Lee, 2005) and analysis of a limited set of mitochondrial data (Zhou et al., 2006). The position of *L. viridis* and therefore Lacertidae was not changed within our MP and NJ topologies. In contrary, a more narrow relationship of Scincomorpha (Lacertidae, Scincidae, Cordylidae) was revealed by these analyses.

Regarding the differences in the phylogenetic pattern of mitochondrial genome sequences compared to nuclear genes and to the morphological phylogeny we strongly recommend that future work is necessary to unravel relationships of the different squamate families. Further studies need to investigate if the differences found in phylogenetic relationships using either mitochondrial genomes or nuclear genes are due to a special feature specific for mitochondrial genomes. Another explanation for the discrepancies between the mitochondrial and nuclear data might be an incomplete taxon sampling of complete mitochondrial genomes within the Squamata. These phylogenetic uncertainties emphasise the need for more complete mitochondrial genome data.

Acknowledgements

We are grateful to Dr. Werner Mayer for kindly providing the isolated mtDNA sample from Austria. Furthermore, we thank Judit Vörös, Dr. Georg Džukić, Katarina Ljubisavljević and Michael Stein for their help to get more samples of *L. v. viridis*. For help with RNA folding we thank P. Stadler and C. Flamm. We thank the anonymous reviewers for the helpful comments on the manuscript and A.D. Sommerfeldt for improving the language of a previous draft. The work of M. U. Böhme was funded by the Evangelisches Studienwerk Villigst e.V. Supplemental material is available on (http://www.bioinf.uni-leipzig.de/IZBI_AG4/SUPPLEMENTS/06_01/).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2007.02.006.

References

Amer, S.A.M., Kumazawa, Y., 2005. Mitochondrial genome of *Pogona vitticepes* (Reptilia; Agamidae): control region duplication and the origin of Australasian agamids. Gene 346, 249–256.

Arevalo, E., Davis, S.K., Sites, J.W., 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Syst. Biol. 43, 387–418.

Böhme, M.U., et al., 2006. Genetic differentiation and diversity of *Lacerta viridis viridis* (Laurenti, 1768) within the northern parts of its range: an investigation using mitochondrial haplotypes. Salamandra 29–40.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.

Brehm, A., Harris, D.J., Alves, C.D., Jesus, J.D., Thomarat, F.D., Vicente, L.D., 2003. Structure and evolution of the mitochondrial DNA complete control region in the lizard *Lacerta dugesii* (Lacertidae, Sauria). J. Mol. Evol. 56, 46–53.

Busack, S.D., Lawson, R., Arjo, W.M., 2005. Mitochondrial DNA, allozymes, morphology and historical biogeography in the *Podarcis vaucheri* (Lacertidae) species complex. Amphib–reptil. 26, 239–256.

Dong, S., Kumazawa, Y., 2005. Complete mitochondrial DNA sequences of six snakes: phylogenetic relationships and molecular evolution of genomic features. J. Mol. Evol. 61, 12–22.

Douglas, D., Janke, A., Arnason, U., 2006. A mitogenetic study on the phylogenetic position of snakes. Zool. Scr. 35, 545–558.

Faith, D.P., Cranston, P.S., 1991. Could a cladogram this short have arisen by chance alone — on permutation tests for cladistic structure. Cladistics 7, 1–28.

Graham, S.W., Olmstead, R.G., Barrett, S.C.H., 2002. Rooting phylogenetic trees with distant ourgroups: a case study from commelinoid monocots. Mol. Biol. Evol. 19, 1769–1781.

Guindon, S., Gascuel, O., 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst. Biol. 52 (5), 696–704.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hedges, S., 1994. Molecular evidence for the origin of birds. PNAS 91, 2621–2624.

Hillis, D.M., Huelsenbeck, J.P., 1992. Signal, noise, and reliability in molecular phylogenetic analyses. J. Heredity 83, 189–195.

Hofacker, I.L., 2003. Vienna RNA secondary structure server. Nucleic Acids Res. 13 (13), 3429–3431.

Janke, A., Arnason, U., 1997. The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent Archosauria (Birds and Crocodiles). Mol. Biol. Evol. 14 (2), 1266–1272.

Janke, A., Erpenbeck, D., Nilsson, M., Arnason, U., 2001. The mitochondrial genomes of the iguana (*Iguana iguana*) and the caiman (*Caiman crocodylus*): implications for amniote phylogeny. Proc. R. Soc. Lond., B Biol. Sci. 268, 623–631.

Kalendar, R., 2005. Fast PCR. Institute of Biotechnology University of Helsinki, Helsinki, Finland.

Kearney, M., Stuart, B.L., 2004. Repeated evolution of limblessness and digging heads in worm lizards revealed by DNA from old bones. Proc. R. Soc. Lond. B 271, 1677–1683.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Kumazawa, Y., 2004. Mitochondrial DNA sequences of five squamates: phylogenetic affiliation of snakes. DNA Res. 11, 137–144.

Kumazawa, Y., 2007. Mitochondrial genomes from major lizard families suggest their phylogenetic relationships and ancient radiations. Gene 388, 19–26.

Kumazawa, Y., Endo, H., 2004. Mitochondrial genome of the Komodo dragon: efficient sequencing method with reptile-oriented primers and novel gene rearrangements. DNA Res. 11, 115–125.

Kumazawa, Y., Nishida, M., 1999. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. Mol. Biol. Evol. 16, 784–792.

Kumazawa, Y., Ota, H., Nishida, M., Ozawa, T., 1998. The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. Genetics 150, 313–329.

Lee, M.S.Y., 2005. Squamate phylogeny, taxon sampling, and data congruence. Org. Divers. Evol. 5, 25–45.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z., Papenfuss, T.J., 1997a. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol. Biol. Evol. 14, 91–104.

Macey, J.R., Larson, A., Ananjeva, N.B., Papenfuss, T.J., 1997b. Evolutionary shifts in three major structural features of the mitochondrial genome among Iguanian lizards. J. Mol. Evol. 44, 660–674.

Macey, J.R., Schulte, J.A., Larson, A., Papenfuss, T.J., 1998. Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement. Mol. Biol. Evol. 15, 71–75.

Macey, J.R., Papenfuss, T.J., Kuehl, J.V., Fourcade, H.M., Boore, J.L., 2004. Phylogenetic relationships among amphisbaenian reptiles based on complete mitochondrial genomic sequences. Mol. Phylogenet. Evol. 33, 22–31.

Macey, J.R., et al., 2005. The complete mitochondrial genome of a gecko and the phylogenetic position of the Middle Eastern *Teratoscincus keyserlingii*. Mol. Phylogenet. Evol. 36, 188–193.

Nylander, J.A.A., 2004. MrModeltest 2.2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Ojala, D., Merkel, C., Gelfand, R., Attardi, G., 1980. The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell 22, 393–403.

Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature 290, 470–474.

Pereira, S.L., 2000. Mitochondrial genome oranization and vertebrate phylogenetics. Genet. Mol. Biol. 23, 745–752.

Rand, D.M., 2001. The units of selection on mitochondrial DNA. Ann. Rev. Ecolog. Syst. 32, 415–448.

Reyes, A., Gissi, C., Pesole, G., Saccone, C., 1998. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Mol. Biol. Evol. 15, 957–966.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sbisa, E., Tanzariello, F., Reyes, A., Pesole, G., Saccone, C., 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene 205, 125–140.

Sokal, R.R., Rohlf, F.J., 1981. Taxonomic congruence in the Leptopodomorpha reexamined. Syst. Zool. 30, 309–325.

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

Tarrio, R., Rodríguez-Trelles, F., Ayala, F.J., 2001. Shared nucleotide composition biases among species and their impact on phylogenetic reconstructions of the Drosophilidae. Mol. Biol. Evol. 18, 1464–1473.

Tavaré, S., 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lect. Math. Life Sci. 17, 57–86.

Townsend, T.M., Larson, A., Louis, E., Macey, J.R., 2004. Molecular phylogentics of Squamata: The position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. Syst. Biol. 53 (5), 735–757.

Uetz, P., 2005. The EMBL Reptile Database. Peter Uetz and the European Molecular Biology Laboratory, Heidelberg, Germany.

Vidal, N., Hedges, S.B., 2004. Molecular evidence for a terrestrial origin of snakes. Proc. R. Soc. Lond. B (suppl.) 271, 226–229.

Vidal, N., Hedges, S.B., 2005. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. C. R Biologies 328, 1000–1008.

Walberg, M.W., Clayton, D.A., 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res. 9, 5411–5421.

Zardoya, R., Meyer, A., 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. PNAS 95, 14226–14231.

Zhou, K.Y., Li, H.D., Han, D.M., Bauer, A.M., Feng, J.Y., 2006. The complete mitochondrial genome of *Gekko gecko* (Reptilia: Gekkonidae) and support for the monophyly of Sauria including Amphisbaenia. Mol. Phylogenet. Evol. 40, 887–892.