Molecular evolution of satellite DNA repeats and speciation of lizards of the genus Darevskia (Sauria: Lacertidae)

Vernata V. Grechko, Doina G. Ciobanu, Ilya S. Darevsky, Sergey A. Kosushkin, and Dmitri A. Kramerov

Abstract: Satellite DNA repeats were studied in Caucasian populations of 18 rock lizard species of the genus Darevskia. Four subfamilies (Caucasian Lacerta satellites (CLsat)I–IV) were identified, which shared 70%–75% sequence similarity. The distribution of CLsat subfamilies among the species was studied. All the species could be divided into at least 3 clades, depending on the content of CLsat subfamilies in each genome: “saxicola”, “rudis”, and “mixta” lizards. CLsatI was found in all studied species, but in very different quantities; the “saxicola” group contained this subfamily predominantly. The “rudis” group also contained CLsatIII, and the “mixta” group carried considerable amounts of CLsatII. The highest concentrations of CLsatI and CLsatII were detected in 2 ground lizards — D. derjugini and D. praticola, respectively. D. parvula predominantly carried CLsatIII. CLsatIV was found only in the Crimean species D. lindholmi. The distribution patterns of satellite subfamilies show possible postglacial speciation within the genus Darevskia. A hybrid origin of species that possess 2 or 3 CLsat subfamilies and important clarifications to the systematics of the genus are proposed.

Key words: tandem DNA repeats, satellites, Darevskia, lizards, reticulate speciation, phylogeography, Caucasus.


Mots clés : ADN répété en tandem, satellites, Darevskia, lézards, spéciation réticulée, phylogéographie, Caucase.

[Traduit par la Rédaction]

Introduction

For a long time, the role of satellites and other noncoding DNA repeats was viewed as a “selfish” (Doolittle and Sapir 1980; Orgel and Crick 1980) or “junk” DNA hypotheses (Ohno 1972). This hypothesis was criticized, on the grounds of biological and philosophical considerations, by Zuckerkandl (1992) and others (reviewed in Nowak 1994; Charlesworth et al. 1994; Comings 1998; Csík and Henikoff 1998; Dimitri and Junakovic 1999; Vernata and Denoeud 2000). In these papers, the authors emphasized the possible role of DNA repeats in the functioning and evolution of living organisms. The remarkable taxon specificity of tandemly organized satellite repeats, and other types of repeats, might reflect this role (reviewed in Elder and Turner 1995; Grechko 2002). Several hypotheses have been put forward that propose a possible general regulatory role for noncoding DNA in the development and evolution of living beings (Trifonov 1999; Korochkin 2002; Boldogkoi 2004).

Satellite repeats are an intrinsic and sometimes predominant part (up to 80%) of all eukaryotic genomes; the number of tandem copies ranges from hundreds to millions in different taxa. Many of the taxa investigated contain more than 1 type of satellite, some of which are more ancient and, therefore, widely distributed among higher taxa (e.g., alpha-like satellites intrinsic to Cetacea (Milinkovitch 1995) and Artio-
dactyla (Blake et al. 1997; Modi et al. 2004)). Other satellite markers are species-specific or genus-specific and can therefore serve to diagnose clades. Genus-specific satellites are considerably different in sequence, length, structure, and organization in various taxa (reviewed in Mielos 1985; Elder and Turner 1995; Grechko 2002). In several instances, satellites, and even microsatellites, are well conserved over very long evolutionary periods, supporting the view that they are subjected to positive selection (Wiedegren et al. 1985; Fitz-Simmons et al. 1995; Pons and Gillespie 2004; Robles et al. 2004). All these facts testify to the biological role of satellites in biodiversity and, hence, in evolution.

However, an apparent lack of data concerning satellites from a large number of living organisms hampers our understanding of their biological role. Some mammals (primate, rodents, ruminantia, cetacea), amphibians, insects, and crustaceans (reviewed in Elder and Turner 1995; Grechko 2002), have been relatively well studied, but very little is known about many other taxa, especially reptiles. There are some data for several lizard genera of fam. Lacertidae (Sauria: Squamata). Monomer satellites of the genus Podarcis (195 bp long), of some European Lacerta (185 bp long) (Capriglione et al. 1998), of the genus Darevskia (146 bp long) (Grechko et al. 1997; Grechko et al. 1998; Roudykh et al. 1999), and of the “Lacerta agilis complex” (160 bp long) (Ciobanu et al. 2004) have been described.

Initially, 2 variants of the satellite from Caucasian rock lizards (genus Darevskia), called Caucasian Lacerta satellite (CLsat), with a sequence difference of about 25%, have been described Roudykh et al. (1999). Since then, the 3rd and 4th variants of CLsat have been found. Hence, the first goal of our study was to conduct a comprehensive analysis of all available data, which meant following the specificities and distribution of all CLsat subfamilies in most of the currently recognized Darevskia species. These data have been analyzed in correlation with zoogeographic aspects and with species genetic similarities, using previously studied molecular markers (Fedorov et al. 1999; Ryabinina et al. 1998; Mayer and Arribas 2003; Carranza et al. 2004). Rock lizards of the Caucasus and Spain have since been separated (Lacerta/Archaeolacerta; they form the genera Darevskia and Iberolacerta, respectively (Arribas 1999). This division is supported by taxonprint data (Grechko et al. 1998), which showed that Caucasian rock lizards from the genus Darevskia form a separate clade as well as the other genera investigated (Podarcis, Eremitta, Ophiosops, Gallotia, Lacerta s. str. and Zootoca).

There are 25 species recognized in the genus Darevskia (7 of which are parthenogenic and not included in this study). In several cases, the species divisions are not obvious and should be verified on the molecular level. In addition, the systematic position of many Darevskia geographic populations and subspecies is ambiguous. Therefore, our second goal was to clarify these taxonomic uncertainties, using satellite markers. Results obtained with several mitochondrial and allozyme markers have been ambiguous and require clarification (Murphy et al. 2000; Fu 2000). We also wanted to determine whether several species of rock lizards originated from interpopulational or interspecific hybridization. It is well known that the parthenogenic Darevskia species has a hybrid origin (Murphy et al. 2000; Ciobanu et al. 2002), and that several bisexual lizard species in the Caucasus freely hybridize (Orlova 1978; Darevsky 1967). The role of interspecies hybridization in animal speciation has been discussed intensively in the past decade (reviewed in Arnold 1997; Avise 2004). The Darevskia complex is a very intriguing group that can be used to study this phenomenon.

Finally, we tested our phylogeographic hypothesis of lacertid distribution in the Caucasus after the last Pleistocene glaciation, using DNA satellites as markers.

Materials and methods

Biological materials

The list of studied species and abbreviations appears in Table 1. Genomic DNA was isolated from the erythrocytes of animals euthanized with Nembutal or chloroform, using a phenol/chloroform extraction of nuclear lysate after proteinase K digestion (Sambrook et al. 1989). DNA concentration was determined from UV absorbance, at 260 nm, and from electrophoresis in agarose gel with DNA standards.

Cloning and sequencing of satellite DNA

Genomic DNA was digested with HindIII or TaqI, separated by electrophoresis in 2% SeaKem agarose gel. DNA fragments of approx 150 bp were ligated into the HindIII or TaqI sites of pGEM-3zf(+) . After transformation into Escherichia coli XL-1 Blue, plasmid DNA was isolated from positive clones selected using the blue–white screening method (Sambrook et al. 1989). Sequencing of plasmid DNA was carried out with dideoxy-chain termination, using kits from Promega (Madison, Wis.) or Sileks M (Moscow, Russia).

DNA hybridization

Hybridization probes were synthesized and labeled with PCR, using specific primers, plasmids carrying particular satellites, and [α-32P]dATP. For CLsatI and for part of CLsatII, we used primers described elsewhere (Ciobanu et al. 2002); we also used 5′-aggtcttcattttagctgatt-3′ and 5′-gaacaacactatc-3′ for CLsatII, and 5′-accccttcatttagctgatt-3′ and 5′-tcaaaacccaagacctegc-3′ for CLsatIII. DNA hybridization was carried out using a Hybond N+ membrane (Amersham), in accordance with the manufacturer’s instructions. All the radioactive probes were hybridized with the same membrane, which contained the same set and quantity of species DNA. Signal intensities were quantified using a Phosphorimager (Packard Instruments) and accompanying OptiQuant Image Analysis software.

Sequence analysis

Homologous nucleotide sequences were analyzed using BLAST (Altschul et al. 1990) and GeneBee similarity
results (Brodskii et al. 1995) in GenBank and EMBL. Pairwise and multiple alignments of satellite DNA were performed using the GeneBee server (http://www.genebee.msu.su) and manually adjusted in the GeneDoc Alignment Editor (http://www.psc.edu/biomed/genedoc). A neighbor-joining tree was constructed using PHYLIP, version 3.63, with default parameters (Felsenstein 1989). Bootstrap support was based on 1000 resampled datasets, using SEQBOOT, PHYLIP, version 3.63.

Accession Nos. of CLsat monomer sequences are as follows: AY262941-9 (chl); gi3087812 (dar1); AY262967-71 (szz1); AY256930-43 (lin1); gi7688051 (val1); gi7688047 (por1); gi7688046 (rad1); gi7688049 (rud1); gi7687990 (alp1); gi7688045 (nai1); gi7688014 (der1); AY2662972-76 (dry2) AY262977-81 (cla1); AY262982-89 (mix1); AY262990-96 (dry1); gi18073593 (cla2); gi7688052 (mix2); gi7687994 (cau2); gi5457400 (dag2); gi5457401 (pra2); gi18073592 (dry3); gi18073595 (mix3); gi18073594 (par3); gi7688016 (lin3); and AY263000-1 (der3).

Results

Satellite repeats CLsat in the genus Darevskia

Satellite monomers initially described in *Darevskia saxicola darevskii* were visible in agarose gel after electrophoresis of DNA digested with HindIII (or TaqI), as the major band of about 150 bp (Grechko et al. 1998; Roudykh et al. 1999). This band was isolated, cloned, and sequenced; the length of monomers was 145 to 147 bp. Repeats were arranged in tandem arrays, as indicated by a 150-bp ladder in Southern hybridization experiments. We named this satellite CLsat.

Genomes of other *Darevskia* species were screened for CLsat. At first, 5 or 6 randomly chosen clones with monomer CLsat inserts isolated from each species were taken for analysis. In the beginning, only CLsat of the *D. s. darevskii* type were detected in *D. s. saxicola, D. s. szczurkii, D. alpina, D. raddii, D. nairiensis,* and *D. chlorogaster* (Table 1). The sequence variability among individual monomers of a species was very low (range, 0%–5%); such individual similarity is typical for other vertebrate satellites (reviewed in Elder and Turner 1995). Variability among species consensus sequences was slightly higher (2%–7%).

More recently, another variant of CLsat, with a total sequence difference of 20% to 25%, was found in several other species. The profound difference between these 2 variants (called CLsatI and CLsatII, respectively) was confirmed after studying the restriction sites for 15 endonucleases; differences in the distribution of 8 restriction sites were revealed (Roudykh et al. 2002). Therefore, the number of randomly chosen clones sequenced was increased so that other possible variants of CLsat were not missed. In several cases, no variant other than CLsatI was found (for instance, in 18 clones of *D. valentini, 10 clones of D. portschinskii,* and 9 clones of *D. chlorogaster*). In several other cases, along with CLsatII, other subfamilies (CLsatIII and CLsatIV) were found.

Figure 1 summarizes these data in the alignment of consensus sequences deduced from the pool of species monomers and from each of the 4 CLsat variants mentioned above. These variants (I, II, III, and IV subfamilies) have sequence similarities of about 75% to 80%. Clones containing CLsatI were found in the genomes of 14 lizard populations considered by zoologists to be systematic species. No other monomer variants were found in 7 of these species (Fig. 1A). DNA hybridization experiments did not detect any significant amount of other CLsat variants in these species. Among CLsatI-containing species, only *D. valentini* and *D. portschinskii* DNA have additionally hybridized with the CLsatIII probe (see below). The clones of the other 5 species included CLsatI and (or) CLsatIII, in addition to CLsatII (Figs. 1B, 1C). CLsatI was revealed only among clones of *D. mixta, D. caucasica, D. daghestanica,* and *D. praticola*; DNA hybridization experiments also showed the presence of CLsatIII and minor CLsatI fractions. Among *D. parvula* clones, there were only CLsatIII satellite sequences, which is the predominant fraction found in hybridization experiments (see below).

CLsatI-containing species are listed in Fig. 1A. As was mentioned above, the variability of CLsatI among species is small (2%–7%), and slightly exceeds intraspecific levels of monomer variability (0%–5%). Diagnostic changes in the alignment of CLsatI distinguish the following 4 subgroups of species: *D. chlorogaster and D. nairiensis* (3 diagnostic substitutions); *D. raddii(G), D. valentini, D. portschinskii,* and *D. radis* (3 diagnostic substitutions); *D. alpina, D. s. darevskii,* and *D. s. szczurkii* (5 diagnostic substitutions); and *D. mixta, D. clarcorum,* and *D. dryada* (3 diagnostic substitutions). Species *D. lindholmi* has no unique sites, whereas *D. derjugini* has 7 autapomorphic substitutions, which sets it apart from other CLsat-containing species. A low number of species-specific changes suggests a very recent divergence of these species (except *D. derju-
Fig. 1. Alignment of specific consensus sequences of Caucasian Lacerta satellite (CLsat) monomers of the *Darevskia* genus. Subfamilies CLsatI (A), CLsatII (B), CLsatIII (C), and CLsatIV (D) are grouped; consensus sequences of each group are shown. Group consensus sequences are aligned in (E). Lower-case letters in consensus sequences designate the variable positions of individual monomers. Fragments conserved in all CLsat monomers are shown as dots and indels are shown as hyphens. Some diagnostic mutations of CLsatI are shown in gray. Short horizontal lines in (E) mark (AT)-rich motifs; gray regions indicate shared nucleotides. For species abbreviations, see Table 1.
Figure 2. Monomer CLsat sequence similarity groups, according to unrooted neighbor-joining tree, based on a comparison of specific consensus sequences. Bootstrap values more than 50% are shown; discussed species subgroups corresponding to CLsat subfamilies are shown in gray. For species abbreviations, see Table 1.

CLsatI

CLsatII

CLsatIII

CLsatIV

Figure 2 shows a neighbor-joining tree, based on consensus sequences represented in Fig. 1, and illustrates the existence of 4 subfamilies of CLsat. Identical topology was obtained with a general analysis that included all individual monomer sequences (data not shown).

Structural properties of CLsat monomer sequences had no peculiar features. As are many similar types of satellites, CLsats are AT-rich (~60%) (Elder and Turner 1995); some conservative AT motifs are shared by all consensus sequences at the same positions (as can be seen in Fig. 1). The relatively high sequence conservation and localization of these motifs in satellites indicates their possible function and, thus, selective constraints. The search for internal similarities in CLsat consensus sequences revealed little similarity (~40%–50%) between the 2 halves of the monomer. It is not improbable that half-sized monomers have given rise to the present CLsat family. Database searches revealed no considerable homologies of CLsat to known nucleotide sequences.

More divergent (rare) structural variants, in addition to the monomers described above, with a difference of approx 5% to 12% (which is below the 25% level of difference be-
tween subfamilies), were found in 7 species from different groups (10% of the total). Such variants are present in *Darevskia* species *nairensis*, *valentini*, *mixta*, *alpina*, *portschinskii*, *caucasica*, and *lindholmi*. Both point mutations and indels were observed. Most rare monomers were found among CLsatIII, and all of them are 3' truncated. For instance, 4 monomers of *D. parvula* were 146 bp long (Fig. 1C), and 10 were truncated to 122–124 bp. The differences corresponding to a fragment of the full-length consensus is within the range of individual variability. 3' terminal truncations of CLsatIII and CLsatIV monomers were also observed in *D. lindholmi*. Two of 40 monomers were truncated (128 and 135 bp), although nothing unusual was observed in their structure.

All 3 available CLsatIII monomers of *D. derjugini* have peculiarities. An 11 bp duplication was revealed in one of them (nucleotide positions 77–98), which increased its length to 158 bp (the sequence in Fig. 1C is shown without this duplication). The first 5' terminal (106 bp long) of this and 2 other unusual monomers (132 bp long) were identical, whereas the 3’-terminal 26 bp fragment of these 2 monomers did not show any similarity to other CLsat monomers. Thus, all unusual CLsatIII variants were either truncated or contained large deletions; more variable monomers of CLsatI and CLsatII had only a number of point mutations. It cannot be ruled out that CLsatIII of this species belongs to a separate, more divergent subfamily, but this remains to be clarified.

**CLsat subfamily distribution in *Darevskia* species, according to DNA hybridization**

Arbitrary selection of DNA clones with the monomers for sequencing represented in Fig. 1 could have led to a loss of minor variants; hence, we verified the presence of all 3 subfamilies with blot hybridization, using the probes for CLsatI (from *D. saxicola darevskii*), CLsatII (from *D. caucasica*), and CLsatIII (from *D. dryada*) monomers. Analysis of the distribution of CLsat subfamilies in the *Darevskia* genus (Fig. 3) allowed us to identify 3 groups of species with higher similarity between them. Although most of the DNA from this species is hybridized with CLsatI more or less intensively, DNA containing predominantly CLsatI (Fig. 3A) is encompassed in the first group. The 2nd group of species demonstrated significant hybridization with both CLsatI and CLsatIII probes; among them, *D. parvula* hybridizes predominantly with CLsatIII probe (Fig. 3B). The 3rd group of species was distinguished by the presence of major CLsatII fraction (Fig. 3C). In addition, all species in this group contained comparable amounts of CLsatI and CLsatIII, except for *D. praticola*, where CLsatII was the major component.

Figure 3 demonstrates considerable variability in the content of each CLsat subfamily among species. The number of CLsat copies in the *D. s. darevskii* has been estimated to be roughly 10⁴ per haploid genome (Roudykh et al. 1999); however, an accurate evaluation of CLsat abundance in all species DNA should be specifically addressed.

The data presented in Figs. 1–3 support the division of the genus *Darevskia* into at least 3 clades; this seems to be in agreement with the data from sequences. We recognize the “saxicola” clade, which is defined by the presence of CLsatI with little sequence heterogeneity as the only major component (Figs. 1 and 3A); the “rudis” clade (Figs. 1 and 3B), which includes species with CLsatIII as the 2nd (or predominant, in the case of *parvula*) component, apart from
CLsatI; and the “mixta” clade, which is distinguished from other clades by the presence of CLsatII, in addition to CLsatI and CLsatIII (Figs. 1 and 3C).

Correlation between geographic distribution of the *Darevskia* species and the distribution of CLsat subfamilies

The pie charts in Fig. 4 represent the content of all CLsat subfamilies in each species, based on the data from Fig. 3, along with the geographic distribution of the studied populations. Most of the lizards studied were caught in locations typical for each species. The species with predominant CLsatI (white sectors in Fig. 4) lies on the vector along the Main Caucasian Range, from southeast (*D. chlorogaster*; Lenkoran, Azerbaijan) to northwest (*D. s. szczerbaki*; Taman Peninsula, Russia). The CLsatI + II + III–containing species (white + striped + black pies in Fig. 4) are arranged on the vector across the Range from southwest (*D. clarcorum*; Turkey) to northeast (*D. daghestanica*; Dagestan, Russia). *D. parvula*, containing predominantly CLsatIII, occurs sympatrically with these species, and on the north Turkish shore of the Black Sea. The CLsatI + III species (black and white pies in Fig. 4) are near the center of the *Darevskia* distributional range.

Discussion

**Lizard satellites**

Our results show that the DNA of lacertid lizards from the genus *Darevskia* contains more than 1 type of satellite, as has been shown for other animals (e.g., Modi et al. 2004). The 4 variants revealed represent 1 satellite family, on the grounds of their conservative parts (at the 75% level) and structural features. Some species of this genus contain 1 predominant type of satellite, as do the “saxicola” group (CLsatI, 7 species), *D. praticola* (CLsatII), *D. parvula* (CLsatIII), and others with the mixture of some or all variants.

These data show that using satellites as phylogenetic markers requires that the possible coexistence of several satellite subfamilies be taken into account; thus, phylogenetic inferences should be drawn only from comparisons with members of the same CLsat subfamily to prevent misleading results. Comparison of CLsatI sequences from *D. saxicola* and *D. mixta* suggests their close relationship, unlike the sequence comparisons of CLsatI of *D. saxicola* and CLsatII or CLsatI of *D. saxicola* and CLsatIII of *D. mixta*. This was also stressed by Baum et al. (2001).

Nevertheless, some phylogenetic and phylogeographic inferences from satellite data can be drawn. First, a supposedly common ancestor of the genus possessed CLsatI, which is shared by all species studied, although in highly variable amounts. Second, the significant degree of similarity within each CLsat subfamily of geographically distant species provides for their fast and directional expansion (Fig. 4). Third, some of the species containing the mixture of different CLsat variants might have originated from interspecific hybridization (discussed below (Was there a hybrid speciation?)).

All these inferences are important for understanding some
Table 2. Arrangement of *Darevskia* species into clades, according to satellite analysis (sat) from this work and mtDNA and allozyme (mt,all) markers (Murphy et al. 2000).

<table>
<thead>
<tr>
<th>“saxicola”</th>
<th>“mixta”</th>
<th>“radis”</th>
</tr>
</thead>
<tbody>
<tr>
<td>sat</td>
<td>mt,all</td>
<td>sat</td>
</tr>
<tr>
<td>dar</td>
<td>dar</td>
<td>mix</td>
</tr>
<tr>
<td>sze</td>
<td>cla</td>
<td>val</td>
</tr>
<tr>
<td>—</td>
<td>bra</td>
<td>dry</td>
</tr>
<tr>
<td>alp</td>
<td>alp</td>
<td>cau</td>
</tr>
<tr>
<td>radL?</td>
<td>—</td>
<td>rad</td>
</tr>
<tr>
<td>nai</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>chl</td>
<td>pra</td>
<td>pra</td>
</tr>
<tr>
<td>der</td>
<td>pra</td>
<td>pra</td>
</tr>
</tbody>
</table>

Note: Species positions supported by all molecular markers are shown in bold. The ambiguous positions of *radG* from Gosh (Armenia) and *radL* from Lenkoran (Azerbaijan) are marked with ? (see below). For species name abbreviations, see Table 1.

of the uncertainties in the taxonomy of *Darevskia* (and Lacertidae in general), which is very complex and is regularly being revised (Arnold 1989; Harris et al. 1998; Arribas 1999; Mayer and Arribas 2003; Carranza et al. 2004). Some attempts in this direction have been undertaken by Capriglione et al. (1998), who isolated 2 unrelated satellites from European lacertids *Podarcis* (195 bp) and *Lacerta graeca* (185 bp), which also differ from *Darevskia* CLsat. The *L. graeca* satellite demonstrated a low similarity (~50%) to the alphoid and CENP-B-like satellites of pri-mates, which led the authors to consider it an ancient alphoid DNA. The 150-bp satellite pSHS from *D. saxicola* was also described (T. Capriglione, personal communication), but proved unrelated to satellite sequences described for the species studied here.

Capriglione et al. (1998) evaluated the genus-level relationships among European lizards, using DNA hybridization with probes constructed on the basis of the above-mentioned satellites. It should be noted that satellite DNA hybridization is not a very reliable approach for evaluating intergeneric similarity, even in closely related genera. In our experiments, 3 related CLsat subfamilies did not cross-hybridize, despite their sequence similarities of ~75% (Fig. 3) (Ciobanu et al. 2003). No hybridization was revealed between *Lacerta agilis* satellites Agi160 and CLsat (about 70% sequence similarity) (Ciobanu et al. 2004). The absence of a DNA hybridization signal cannot unambiguously confirm the absence of similar satellites in studied taxa. Therefore, the presence of CLsat-like repeats in *Zootoca, Podarcis, Eremias, Ophisops,* and *Gallotia* genomes (Ciobanu et al. 2003) cannot be ruled out on the basis of only hybridization data; sequence data are necessary.

**Phylogeography of the *Darevskia* species**

The phylogeographic approach, developed and proven to be useful by Avise (2004), permits the investigation of the “dispersal of taxa through a region, speciation, adaptive radiation, and extinction; in other words, investigation of the fundamental links between population process and regional patterns of diversity and biogeography.” In this context “the direct determination of DNA nucleotide sequences has permitted...fruitful cross-taxa comparisons in evolutionary history” (Bermingham and Moritz 1998). Our results seem to contribute to this field.

Our data show that the variation in the content of CLsat subfamilies in different species delineates at least 3 clades (Table 2, Fig. 3). The geographic positions of species of the “saxicola” and “mixta” clades (Fig. 4) might reflect the direction of pathways (vector) of the rock lizard’s secondary (postglacial) speciation, along their migration pathways from the Asia Minor or Transcaucasian refugia. Concepts of the last wave of speciation in Europe, following the Pleistocene glaciation, are being formulated (reviewed in Taberlet and Cheddadi 2002; Hewitt 2004). The Caucasus is not being specifically discussed in these reconstructions, but there is no reason to exclude it from such analysis. It is well known that Pleistocene glaciation expanded to the Cau-casian isthmus. According to Berg (1952), glaciation covered the entire Main Caucasian Range and some side ridges, and the ice level during this period fell from 4000–5000 m to 1500 m, or less, above sea level. Together with the cooling of the climate, this resulted in mass extinction (or southward extrusion) of the fauna. About 10 000 to 15 000 years ago, this region was released from the so-called Bezengi glaciation (somewhat contemporary to Wurm glaciation in the European part of modern Russia). Along with climatic warming, the ridges and gorges were released from ice, and the current glaciation level rose to 3000–4000 m. At least 2 low refugia were not exposed to glaciation — the Colchis Lowland in the extreme southwest of Transcausasia and the Lenkoran Lowland in southeastern Azerbaijan. The maintenance of the Tertiary fauna and flora in these refugia suggests that ancestors of current lacertids could have survived there before recolonization of the Caucasus. Although almost no palaeontological records are available for ancestors of Caucasian lacertids, it is possible that the contemporary species inhabiting the refugia resemble them most, particularly *D. parvula* and *D. clarrorum/D. dryada* in the southwest and *D. chlorogaster* in the southeast (Darevsky 1967).

The length of the “saxicola” vector is at least twice that of the mixta vector crossing the Main Caucasian Range, and it is quite interesting that species at the start and end points of this vector are practically identical in sequences and uniqueness of CLsat (Fig. 1A). Their movement along a tentative vector was probably much faster than the movement of the “mixta” group across the Caucasian ridge. The distribution of CLsatII-containing species along the vector crossing the ridge through the high passes demonstrates that they could have overcome this natural barrier; *D. caucasica,* which largely inhabits the northern slope, also inhabits the southern slopes of the Krestovyi and Roksky Passes (~3000 m above sea level), i.e., along this pathway.

**Was there a hybrid speciation?**

This analysis correlates well with the “pioneer-phalanx” hypothesis of Nichols and Hewitt (reviewed in Hewitt 2004). A “pioneer” lineage of the “saxicola” clade occu-
occupied a vast territory during a short evolutionary period after glaciation, and was dispersed mainly along the Caspian shore and the northern slope of the Caucasian ridge. This more rapidly moving lineage has maintained only 1 major variant of CLsat. The remainder of the species moved along the southern slope toward Georgia and the Armenian Up- land, from both Black and Caspian refugia, to form the “phalanx”, accumulated in the very small territory, and made up the lineages combining several CLsat subfamilies.

This hypothesis raises the possibility that natural interspecific crossing is a source of gene flow in complex “phalanx” populations: the question of why some lizard species possess single major satellite variants while others have 2 or more major variants remains. The hybrids could have retained 2 or 3 major CLsat subfamilies from ancestral parental populations. It is known that interspecific hybridization can promote large-scale chromosome rearrangements, a high frequency of recombination, and other events that accom- company the speciation process and (or) take part in it (reviewed in Fontdevila 2005). The problem of the possible hybrid origin of many species has been intensively discussed recently, and is supported by examples from both the plant and animal kingdoms (reviewed in Bullini 1994; Arnold 1997; Dowling and Secor 1997; Barton 2001; Avise 2004).

Our results can also be explained in terms of the popular “satellite library” hypothesis (reviewed in Micos 1985), which allows for the explosive amplification of certain satellite monomers. The mutant forms of the ancestral “master” gene could have amplified differently in various lineages, forming different patterns of satellite subfamilies.

Numerous observations favor the hypothesis of hybridogenic-driven speciation in a phalanx community of rock lizards. It is well known that populations of Caucasian lizards from the genus Darevskia can freely hybridize at their range boundaries. This is true for viable and persistent parthenogene- netic lizard species, the hybrid origin of which was predicted and confirmed by the distribution of parental CLsat variants in the parthenogenetic Darevskia species (Ciobanu et al. 2002). Apart from parthenogenetic forms, Darevsky (1967) reported hybridization between a number of Cauc- asian bisexual lizard species. The hybrid origin of D. mixta from a cross between D. derjugini and D. saxicola was ini- tially proposed, although this proposal was later questioned (Uzzell and Darevsky 1975). The satellite data presented here revive this problem, assuming the involvement of D. derjugini and some CLsatII- or CLsatIII-containing ancestors (e.g., D. parvula instead of D. saxicola) in the hy- bridogenesis. Such hybrids are reported from natural popula- tions in the sympatric zone in Georgia (Darevsky 1967).

Viable and fertile natural or experimental hybrids combin- ing the morphological characteristics of both parents have been reported for D. clarcorum × D. rudis, D. dryada × D. rudis, and for D. derjugini with D. parvula, D. mixta, D. saxicola, and D. caucasia (Orlova 1978). As Orlova (1978) mentioned, modern populations are morphologically highly variable, and 1 or both parental forms can be com- pletely substituted by their hybrids, as demonstrated by D. derjugini × D. parvula hybridization. These data corre- late with our hypothesis based on satellite analysis.

There are at least 3 species of Caucasian lacertids that can be considered to be progenitors in potential reticulate speci-

ation. The ground lizard D. derjugini, which contains more divergent CLsat as a major satellite (Fig. 1A), is among them. Other possible progenitors include the ground lizard D. praticola and the above-mentioned D. parvula, which contain predominantly CLsatII and CLsatIII, respectively, as major satellites.

Molecular data in verification of Darevskia taxonomy, based only on morphological considerations

All studied species share the presence of CLsatI in their genomes, which does not contradict the taxonomic isolation of Darevskia from the large and diverse genus Lacerta. The recognition of 3 clades within this genus based on CLsat variants, as proposed here, largely agrees with their subdivi- sion based on mtDNA and allozyme analysis (Murphy et al. 2000). All molecular markers reveal the same 3 clades in Darevskia, and the positions of 13 species are the same (Table 2). As would be expected, the 2 subspecies of D. sax- icola — D. s. darevskii and D. s. szczerbaki — studied here, and D. s. darevskii and D. s. braunery studied by Murphy et al. (2000) fall into the “saxicola” clade. The 4th subspecies, D. s. saxicola (not studied here), which is highly similar to these 3 subspecies in both morphological and random ampli- fied polymorphic DNA markers (Ryabinina et al. 1998), might also belong to this clade. Murphy et al. (2000) as- signed D. parvula to the “rudis” clade, which is in agree- ment with our data on the structure of CLsatIII. At the same time, the presence of CLsatIII in the “mixta” clade calls into question the assignment of D. parvula to the “rudis” clade. Our data place D. parvula phylogenetically outside these 2 clades, and suggest that a common ancestor of the clades have hybridized with D. parvula.

Our data contradict the assignment by Murphy et al. (2000) of D. derjugini to the “mixta-caucasica” clade (Table 2) and the assignment of D. praticola to the “sax- icola” clade. Although Murphy et al. (2000) proposed it on the basis of mtDNA and allozyme analysis, they admitted that support for this grouping was ambiguous. We believe that the allozyme (although Murphy et al. consider them less significant) and our satellite data (the presence of CLsatI) confirm that D. derjugini is the sister taxon to a clade comprising all other studied Darevskia species. The position of D. praticola remains ambiguous; its assignment to the “saxicola” clade (Murphy et al. 2000) is not con- firmed by the presence of CLsatII (synapomorphic for D. praticola and the “mixta” clade), and the “saxicola” clade lacks it (Fig. 3).

The origin of the Crimean D. lindholmi remains a mys- tery; the available data do not exclude the hybrid origin of this species, which contains all 4 CLsat subfamilies. With these features, D. lindholmi shows affinities to the D. mixta group, which is geographically more distant than, for example, the CLsat-containing D. s. szczerbaki. At the same time, D. lindholmi DNA includes an autapomorphic feature — a unique variant of CLsat (CLsatIV), which has been found in no other Darevskia species (Fig. 2).

Finally, the positions of D. raddei and D. nairensis cannot be considered definitive without a detailed investigation of a number of populations, called the “raddei complex”. This is confirmed by the differences revealed for this species from Armenia (radG in Table 1) and Azerbaijan (radL in Ta-
ble 1). The population from Gosh is sympatric with some species of the "rudis" clade, and has the same set of satell­
ites (CLSatI + CLSatIII). The population from Lenkoran is sympatric with D. chlorogaster and, similarly, has CLSat alone. That is why we preliminarily assigned D. raddei pop­
ulations to both clades; however, the situation could be more complicated because Murphy et al. (2000) assigned it to an­
other clade — "mixta/caucasia" — on the basis of mtDNA and allozyme analysis (Table 2).

Thus, the systematics of the genus Darevskia is far from being a closed case. We hope that the results presented here and in other papers dedicated to the molecular basis of the systematics of this taxon will provide some new ideas on the taxonomic ranging of these reptiles.

Acknowledgements

We thank F. Danielyan, B. Tuniyev, K. Milto, D. Ryabinin, N. Ryabinina, I. Martirosyan, I. Rudykh, N. Vassetzky, V. Orlova, M. Arakelyan, and A. Ryskov for their help in ani­mal collection and for discussions of the results. This work was supported by grants of the Russian Foundation for Basic Research (Nos. 03-04-49157, 05-04-48147, and 06-04-48222) and of the Russian Project “Biodiversity and Dynamics of Gene Pools” (subprogram II — “Dynamics of Gene Pools”). We are indebted to Brian Golding and to 2 anonymous reviewers for valuable comments and correc­
tions.

References


Carranza, S., Arnold, E.N., and Amat, F. 2004. DNA phylogeny of lacerta (Iberolacerta) and other lacertine lizards (Reptilia: Laceri­tidae); did competition cause long-term restriction? Systematics and Biodiversity, 2: 57–77. doi:10.1017/S147720000401355.


Darevsky, I.S. 1967. Skal’nye yashzeritsy Kavkasa (Caucasian rock lizards). Nauka, Leningrad. [In Russian.]


