ISSN 0026-8933, Molecular Biology, 2006, Vol. 40, No. 2, pp. 333–336. © Pleiades Publishing, Inc., 2006.
Original Russian Text © S.A. Kosushkin, O.R. Borodulina, V.V. Grechko, D.A. Kramerov, 2006, published in Molekulyarnaya Biologiya, 2006, Vol. 40, No. 2, pp. 378–382.

SHORT COMMUNICATIONS

UDC 577.21

New Family of Interspersed Repeats from Squamate Reptiles

S. A. Kosushkin, O. R. Borodulina, V. V. Grechko, and D. A. Kramerov

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia; e-mail: grechko@eimb.ru

Received November 23, 2005

DOI: 10.1134/S0026893306020221

Key words: SINE, retroposon, reptiles

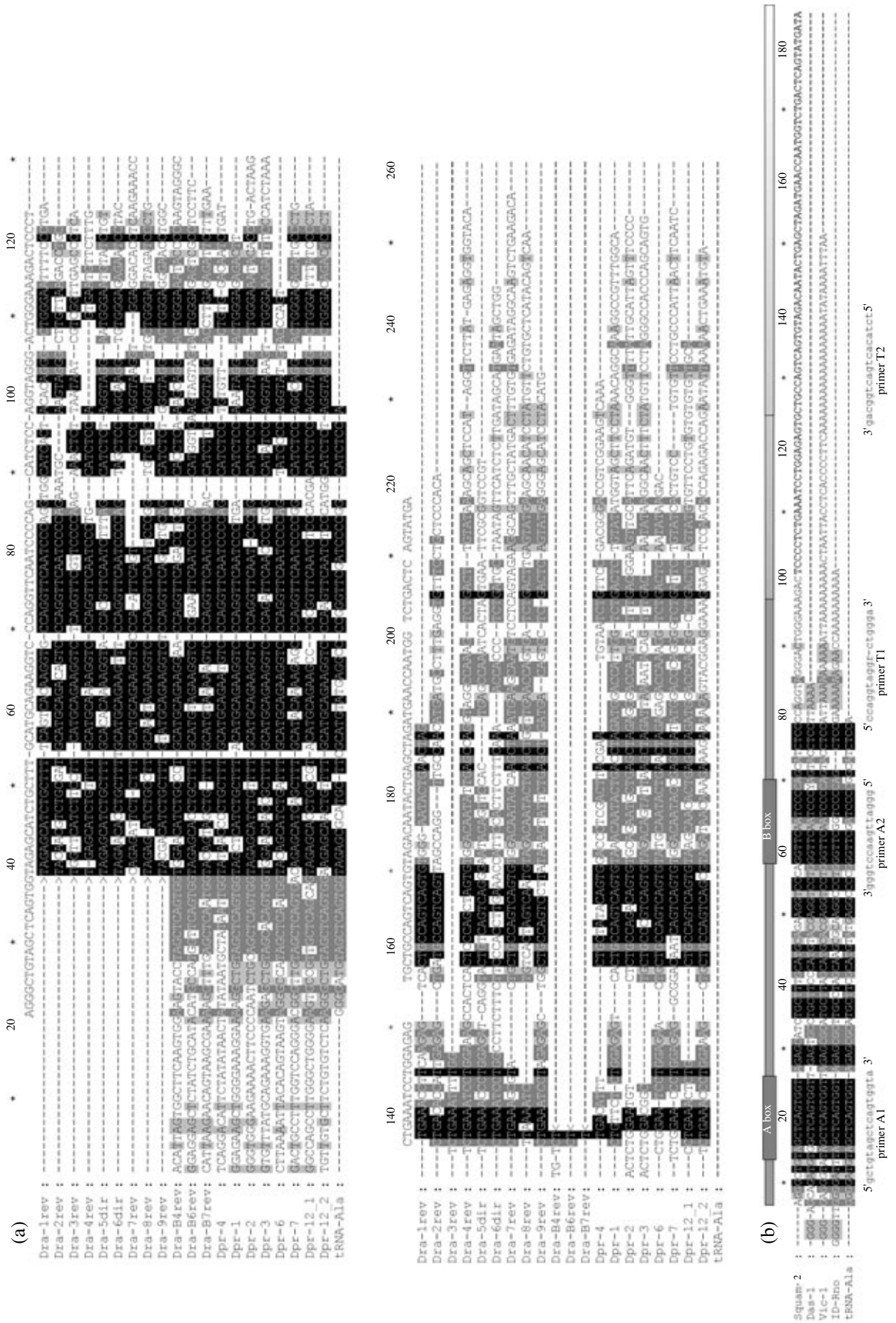
Short interspersed elements (SINEs) account for a considerable portion of the genome in many vertebrates. Mammalian and fish SINEs have been studied most extensively, making it possible to elucidate the main features of SINE structure and evolution. The 5' end of SINE is usually tRNA-related and is often followed by a specific sequence, the 3' part of which is similar to the 3'-terminal region of a long interspersed element (LINE) from the genome of the same species [1]. The 3' end of SINE has either an array of short direct tandem repeats or oligo(A), which are presumably necessary for reverse transcription and integration of the DNA copy into the genome [2, 3].

Until recently, little was known on the SINEs of squamate reptiles (lizards and snakes, the order Squamata). Only one such repeat has been found in the genome of lizard *Podarcis sicula* (fam. Lacertidae); its 5'-terminal region is similar to $tRNA^{Lys}$ and the 3'-terminal region, to a LINE-related element [4]. However, only a single copy of this element has been described and there are no data on its distribution in other reptilian genomes.

Here we describe a new family of SINE-related interspersed repeats from two lizards of the genus *Darevskia* (Lacertidae): *D. praticola* and *D. raddei.* Differences between their copies do not exceed the level of intraspecific divergence (Fig. 1). The repeat was found as a result of sequencing of PCR-amplified fragments from genomic DNA, as described previously [5, 6]. The primer was complementary to a region containing the A box of the RNA polymerase III promoter (A-PCR). The nine sequenced amplification products included DNA fragments that contained the 3' SINE region starting with the A box (5'-TRGCTCAGTGG-3', indicated with > in Fig. 1a). Based on their sequences, we constructed a reverse primer, 5'-GCTCTCYAGGRTTTCA-3' (indicated with < in Fig. 1a), which allowed us to isolate three additional sequences containing the 5' region of SINE. In addition, eight full-length copies were obtained by screening a *D. praticola* genomic library with a labeled probe. The probe was amplified with primers A1 and T2 directed to conserved internal sequences of SINE (Fig. 1b).

Analysis showed that all sequences have structural elements characteristic of canonical SINEs (Fig. 1a). The 5' region displays a considerable similarity to mammalian $tRNA^{Ala(CGC)}$ (Fig. 1a), which contains elements of the internal promoter for RNA polymerase III (boxes A and B). This region is followed by a tRNA-unrelated sequence of about 100 bp, which is specific for reptiles of the order Squamata. The tRNA-related region (85 bp) is most conserved among all copies; conservation involves the 10–15 bp located immediately downstream of the tRNA-related region. This sequence is followed by a short, more variable region. The remaining part varies unevenly along its length. For instance, the conserved region is followed by a 15-bp sequence with low similarity; further downstream, the divergence slightly decreases and then increases. In total, the tRNA-unrelated region is insignificantly longer than the corresponding tRNA.

At the same time, the repeat has some features that distinguish it from the canonical SINEs of mammals and fish [1, 7, 8]. For instance, we did not detect a poly(A) tail or short direct repeats at the 3' end of the element. SINEs lacking a poly(A) tail (so-called tailless retroposons) have been found in most mammalian genomes and probably utilize other mechanisms for interacting with reverse transcriptase during retroposition [9]. In addition, the sequences adjacent to the new element lack flanking short direct repeats, which are duplications of the putative target site for insertion of new SINE copies. Hence, we could not exactly demarcate the element and establish its length, which is about 160–180 bp. It is possible to assume that elements of the new family lost retroposition-associated regions, namely, short tandem repeats at the 3' end of an element and the flanking direct repeats, as characteristic of other ancient SINE families (e.g., MIR [10]).


(a)
(b)
A box
B box
primer A1
primer A2
primer T1
primer T2

Alternatively, we could not exclude that the retroposition mechanism utilized by reptilian SINEs differs from that of other SINEs, which explains the lack of the above structural elements.

To check whether the new SINE occurs in DNA of other taxa, dot blot hybridization with a specific probe was carried out with genomic DNAs of representatives of 12 lizard families of the order Squamata, two snake families, tortoises, crocodiles, amphibians (brown frog), mammals (mouse and human), and three orders of birds (table). Hybridization was performed separately with two labeled probes. One was obtained with primers A1 and A2 and was complementary to the most conserved (tRNA-related) region of the repeat. The other was amplified with primers T1 and T2 and was complementary to the tRNA-unrelated region, which is specific for squamate reptiles.

With both probes, hybridization signals were detected from DNAs of all squamate reptiles, though weak in the case of chameleon, iguana, and agama from the clade Iguania, and the two snakes, *Boa constrictor* and *Elaphe dione* (table). DNAs of all other animals, belonging to taxa other than the order Squamata, produced no hybridization signal with either primer, or showed weak hybridization with the primer directed to tRNA. Based on these findings, we termed the new interspersed repeat Squam-2.

Human, mouse, and frog DNAs produced weak hybridization signals with the probe corresponding to the $tRNA^{Ala(CGC)}$-related region. The similarity with $tRNA^{Ala(CGC)}$ is characterized in Fig. 1b. A possible explanation is that the rodent genome contains elements of the SINE ID family, which also originates from $tRNA^{Ala}$. The signals from human and frog DNAs could be attributed to tRNA genes or pseudogenes present in the corresponding genomes. However, hybridization with the tRNA-unrelated probe was not observed for these species, suggesting the lack of Squam-2 from their DNAs.

To verify the results of hybridization, DNAs of all animals were tested by PCR with primers A1 and T2. The reaction product corresponding in size to SINE was obtained for all but one representative (chameleon) of the order Squamata. The lack of the specific PCR product in the case of chameleon correlates with weak hybridization of its DNA (table). This issue needs further investigation. It is noteworthy that the

Distribution of Squam-2 in the genomes of animals from various taxa as revealed by hybridization and PCR

Taxonomic position			Hybridization		PCR
			A1 + A2	T1 + T2	
Mammalia: Eutheria	Hominidae	1	±	–	±
	Rodentia	2	±	–	–
Mammalia: Marsupialia	Macropodidae	3	–	–	–
Aves	Strigidae	4	–	–	–
	Falconidae	5	–	–	–
	Passeridae	6	–	–	–
Amphibia: Anura	Ranidae	7	+	–	–
Reptilia: Squamata	Chamaeleonidae	8	+	±	–
	Agamidae	9	+	+	+
	Iguanidae	10	+	+	+
	Lacertidae	11	+	+	+
	Lacertidae	12	+	+	+
	Cordylidae	13	+	+	+
	Scincidae	14	+	+	+
	Varanidae	15	+	+	+
	Anguidae	16	+	+	+
	Teiidae	17	+	+	±
	Helodermatidae	18	+	+	+
	Eublepharidae	19	+	+	+
	Gekkonidae	20	+	+	+
	Colubridae	21	±	±	±
	Boidae	22	±	+	+
Reptilia: Testudines	Testudinidae	23	–	–	–
Reptilia: Crocodylia	Crocodylidae	24	–	–	–

Note: DNAs were hybridized with the probes amplified with primers A1 + A2 or T1 + T2 (Fig. 1). Designations: (+), an intense hybridization signal and a major band of the PCR product; (–), no hybridization signal and no PCR product; (±), a weak hybridization signal and a minor band of the PCR product. The following species were examined: 1, *Homo sapiens*; 2, *Mus musculus*; 3, *Dendrolagus bennettianus*; 4, *Asio otus*; 5, *Falco tinnunculus*; 6, *Passer domesticus*; 7, *Rana temporaria*; 8, *Chamaeleo calyptratus*; 9, *Chlamidosaurus kingii*; 10, *Iguana iguana*; 11, *Darevskia raddei/praticola*; 12, *Gallotia* sp.; 13, *Cordylus giganteus*; 14, *Tiliqua* sp.; 15, *Varanus prasinus*; 16, *Anguis fragilis*; 17, *Tupinambis teguixin*; 18, *Heloderma suspectum*; 19, *Eublepharis macularius*; 20, *Gekko gecko*; 21, *Elaphe dione*; 22, *Boa constrictor*; 23, *Testudo graeca*; and 24, *Crocodylus niloticus*.

Fig. 1. (a) Squam-2 sequences of lizards of the family Lacertidae: Dra, *Darevskia raddei* and Dpr, *D. praticola.* The nine upper sequences were obtained by PCR with the primer directed to the A box (indicated with >); the next three sequences were obtained with the reverse primer (<). The eight lower full-length sequences were obtained by sequencing cloned fragments from a genome library; tRNA-Ala is the $tRNA^{Ala(CGC)}$ sequence from the mouse and human genomes (identical sequences were found in many vertebrate genomes by a computer search). (b) Alignment of the $tRNA^{Ala(CGC)}$-related consensus sequences of the SINEs Squam-2 of lizards, Das-1 of armadillo [6, 14], Vic-1 of vicuna [13], and ID-Rno of rodents [12]. Primers A1, A2, T1, and T2 are shown at the bottom. A scheme of Squam-2 (at the top) reflects a decrease in sequence conservation towards the 3' end and shows the A and B boxes of the promoters for RNA polymerase III.

two snakes from different families differ in the presence of Squam-2, as observed with both methods. The *Boa constrictor* genome contains more copies of the element as compared with the *Elaphe dione* genome, where the copies are more divergent or fewer. These findings make it possible to assume that the SINE under study arose in the genome of an ancestor of squamate reptiles about 250 Myr ago, according to the paleontological dating of the order Squamata [11]. After that, SINEs of the family were spread and diverged to a various extent in the genomes of different families of the order.

Using the BLAST program, we searched GenBank for sequences similar to the tRNA-unrelated part of SINE and found no homologs. Similarity with the $tRNA^{Ala(CGC)}$-related region was observed only for mammalian SINEs, originating from the same tRNA. In particular, these were ID of rodents [12], Vic-1 of vicuna [13], and Das-1 of armadillos [6, 14] (Fig. 1b). Possibly, $tRNA^{Ala}$ gave origin more than once in evolution to taxon-specific families of SINEs with a relatively simple structure.

To summarize, we found and characterized a new SINE, Squam-2, in DNAs of reptiles of the order Squamata. Squam-2 belongs to the repeats originating from $tRNA^{Ala(CGC)}$ and known in mammals. Yet some properties distinguish Squam-2 from SINEs with the canonical structure, established with elements from representatives of other classes. Individual copies of Squam-2 are greatly divergent in the tRNA-unrelated region and virtually lack a common 3'-terminal structure. Simultaneously, we found another SINE family, Squam-1, in squamate reptiles. Data on this family have been prepared for publication.

ACKNOWLEDGMENTS

We are grateful to D.B. Vasil'ev and I.S. Darevsky for advice and blood samples of several animals.

This work was supported by the Russian Foundation for Basic Research (project nos. 03-04-49157, 05-04-49553).

REFERENCES

1. Ohshima K., Hamada M., Terai Y., et al. 1996. The 3' ends of tRNA-derived short interspersed repetitive elements are derived from the 3' ends of long interspersed repetitive elements. *Mol. Cell Biol*. **16**, 3756–3764.
2. Okada N. 1991. SINEs. *Curr. Opin. Genet. Dev.* **1**, 498–504.
3. Schmid C.W. 1996. Alu: Structure, origin, evolution, significance, and function of one-tenth of human DNA. *Prog. Nucleic Acid Res. Mol. Biol*. **53**, 283–319.
4. Fantaccione S., Russo C., Palomba P., et al. 2004. A new pair of CR1-like LINE and tRNA-derived SINE elements in *Podarcis sicula* genome. *Gene*. **339**, 189–198.
5. Borodulina O.R., Kramerov D.A. 1999. Wide distribution of short interspersed elements among eukaryotic genomes. *FEBS Lett*. **457**, 409–413.
6. Borodulina O.R., Kramerov D.A. 2005. PCR-based approach to SINE isolation: Simple and complex SINEs. *Gene*. **349**, 197–205.
7. Kajikawa M., Okada N. 2002. LINEs mobilize SINEs in the eel through a shared 3' sequence. *Cell*. **111**, 433–444.
8. Ohshima K., Okada N. 2005. SINEs and LINEs: Symbionts of eukaryotic genomes with a common tail. *Cytogenet. Genome Res*. **110**, 475–490.
9. Schmitz J., Churakov G., Zischler H., et al. 2004. A novel class of mammalian-specific tailless retropseudogenes. *Genome Res*. **14**, 1911–1915.
10. Jurka J., Zietkiewicz E., Labuda D. 1995. Ubiquitous mammalian-wide interspersed repeats (MIRs) are molecular fossils from the mesozoic era. *Nucleic Acids Res*. **23**, 170–175.
11. Carroll R.L. 1987. *Vertebrate Paleontology and Evolution*. New York: Freeman.
12. Kim J., Martignetti J.A., Shen M.R., et al. 1994. Rodent BC1 RNA gene as a master gene for ID element amplification. *Proc. Natl. Acad. Sci. USA*. **91**, 3607–3611.
13. Lin Z., Nomura O., Hayashi T., et al. 2001. Characterization of a SINE species from vicuna and its distribution in animal species including the family Camelidae. *Mammal. Genome*. **12**, 305–308.
14. Churakov G., Smit A.F., Brosius J., et al. 2005. A novel abundant family of retroposed elements (DAS-SINEs) in the nine-banded armadillo (*Dasypus novemcinctus*). *Mol. Biol. Evol*. **22**, 886–893.