ISSN 0026-8933, Molecular Biology, 2007, Vol. 41, No. 6, pp. 986–989. © Pleiades Publishing, Inc., 2007.
Original Russian Text © N.A. Veniaminova, K.P. Gogolevsky, N.S. Vassetzky, D.A. Kramerov, 2007, published in Molekulyarnaya Biologiya, 2007, Vol. 41, No. 6, pp. 1081–1084.

SHORT COMMUNICATIONS

UDC 577.21

Priority publication

Comparative Analysis of the Copy Number of ID and B1 Short Retroposons in Rodent Genomes

N. A. Veniaminova, K. P. Gogolevsky, N. S. Vassetzky, and D. A. Kramerov

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

Received April 25, 2007

Accepted for publication May 22, 2007

DOI: 10.1134/S0026893307060155

Key words: SINE, retroposon, DNA dot hybridization, rodents, Rodentia, genomics, evolution

The mobile genetic elements known as short retroposons, or SINEs (Short Interspersed Elements), are repeated sequences with a length of 80 to 400 bp that are dispersed over the eukaryotic genome and are amplified via reverse transcription [1]. The genome of a mammalian species usually contains two–four families of short retroposons, each composed of tens to hundreds of thousands of copies. As a rule, their nucleotide sequences display a 65–90% similarity. SINEs are transcribed by RNA polymerase III owing to the presence of the type II RNA polymerase III promoter in their 5'-terminal region. The promoter consists of two boxes (A and B) separated by 30–40 bp. Short retroposons belong to nonautonomous mobile elements, as they do not encode their own enzymes and utilize for their reproduction the reverse transcriptase encoded by a long retroposon, or LINE (Long Interspersed Element). The majority of mammalian SINEs reproduce with the help of long retroposon L1.

The SINE families originating from tRNA molecules are the most widespread, as their nucleotide sequences are similar to certain tRNAs. Two families of short retroposons—B1 of rodents and *Alu* of primates [2–5]—originate from the cytoplasmic 7SL RNA [6, 7], which is contained in ribonucleoprotein signal recognition particles (SRPs). SRPs are involved in recognition of the signal peptides of secreted and membrane proteins.

We have recently demonstrated that B1 elements are widespread in rodents, which form the largest and most diverse mammalian order, comprising over 30 families [8, 9]. Considerably less is known about the abundance of the ID element, another rodent SINE [10]. This element has been well studied in the rat, mouse, and guinea pig genomes [11]. ID originates from alanine $tRNA^{CGC}$. One of the ID copies became a gene for the small untranslated BC1 RNA. This RNA, specific to rodents [12], is involved in mRNA translation regulation in neurons [13]. In turn, the BC1 RNA became a direct evolutionary ancestor of an ID subfamily, which then gave origin to other subfamilies [11]. The ID element is about 100 bp, contains a tRNA-related region and an A-rich tail, and belongs to the so-called simple SINEs [14].

In this work, we studied the abundance of ID in rodents belonging to 21 families and demonstrated that this repetitive DNA sequences is present in all species analyzed. The ID and B1 copy numbers were compared for various rodent genomes. Dimeric SINEs composed of these two elements were discovered.

Figure 1 shows the results of dot hybridization of genomic DNAs from 29 rodent species with ^{32}P-labeled ID and B1 elements of the common house mouse *Mus musculus*. DNAs of all rodent species gave hybridization signals; however, their intensities varied in a wide range. The radioactivity of each dot was determined with a Cyclon Phosphorimager; the ID and B1 copy numbers in the rodent genomes were assessed based on the ID copy number in the *Rattus norvegicus* genome (225,000) and the B1 copy number in the *M. musculus* genome (564,000) [15, 16]. The majority of rodent families

in this experiment were represented by one species except for Muridae (mice, rats, and gerbils), Dipodidae (jerboas), and Sciuridae (squirrels), each represented by several species.

Figure 2 shows the histogram of ID and B1 copy numbers in the genomes of various rodent families. The copy numbers varied from 8000 (*Tbo,* pocket gopher, Geomyidae) to 650,000 (*Mso,* social vole, Cricetidae) for B1 and from 25,000 (*Mmu*, house mouse, Muridae) to 500,000 (*Tas*, Asian chipmunk, Sciuridae) for ID. As a rule, the copy number of B1 was higher than that of ID in the rodents belonging to Myodonta (Muridae, Cricetidae, Spalacidae, Rhizomyidae, Zapodidae, and Dipodidae); the representatives of the group Gliridae–Sciuridae–Aplodontidae displayed an opposite pattern. This suggests considerable distinctions in the retroposition activities of B1 and ID during the evolution of different rodent families.

A screening of several rodent genomic libraries with the *M. musculus* ID probe and subsequent sequencing of the hybridizing fragments did reveal the ID nucleotide sequences in all rodent families studied. In addition to common single ID elements, we discovered SINEs composed of two parts (monomers) in many rodent species. In some of them, ID was the first (left) monomer and B1 was the second (right), whereas other species displayed the inverse arrangement of the monomers. In certain cases, the ID sequence in dimeric SINEs contained specific deletions or small additional regions. We have earlier described several dimeric SINE families, namely, MEN in the ground squirrel *Menetes berdmorei* [17], B1-dID in dormice (Gliridae) and squirrels (Sciuridae) [18, 19], and IDL-Geo in gophers (Geomyidae) and pocket mice (Heteromyidae) [20]. The last SINE contains a sequence of an unknown origin, Geo, as the right monomer instead of B1. Presumably, the presence of IDL-Geo in combination with a small B1 copy number determines a uniquely high ratio (50 : 1) of ID to B1 sequences in the pocket gopher genome (Fig. 2, *Tbo*). The dimeric SINEs discovered in this work in the genomes of rodents from the families Pedetidae, Anomaluridae, Castoridae, Thryonomyidae, Myocastoridae, and several others will be described in detail elsewhere.

The results demonstrate that, first, ID and B1 are characteristic of the entire order of rodents; second, the retroposition rate of these elements essentially differs among the genomes of various rodent species;

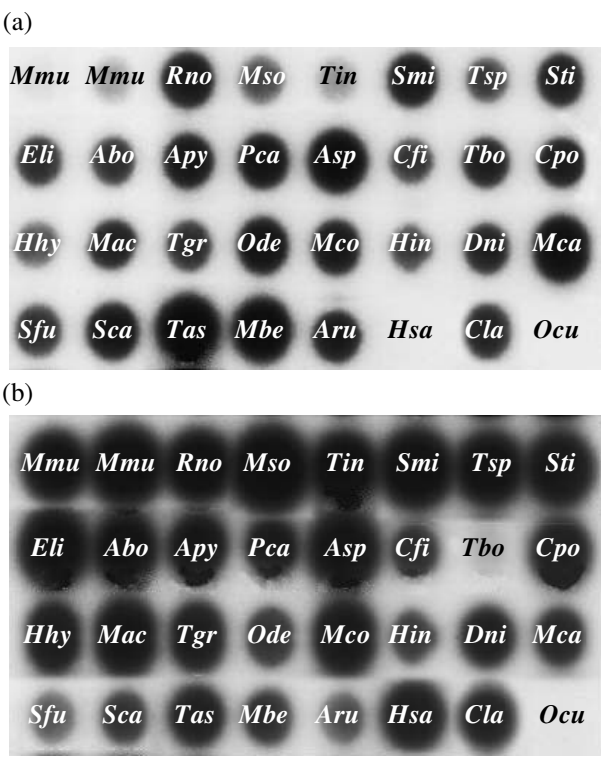

Fig. 1. Dot hybridization of genomic DNA of various rodent species with the mouse (a) ID and (b) B1 elements as probes. Here and in Fig. 2: Murinae: *Mmu*, *Mus musculus* (house mouse) and *Rno*, *Rattus norvegicus* (brown rat). Cricetidae: *Mso*, *Microtus socialis* (social vole). Gerbilinae: *Tin*, *Tatera indica* (large naked-soled gerbil). Spalacidae: *Smi*, *Spalax microphthalmus* (Russian mole rat). Rhizomyidae: *Tsp*, *Tachyoryctes splendens* (East African mole rat). Zapodidae: *Sti*, *Sicista tianshanica* (Tien Shan birch mouse). Dipodidae: *Eli*, *Eremodipus lichtensteini* (Lichnetstein's jerboa); *Abo*, *Allactodipus bobrinskii* (Bobribski's jerboa); and *Apy*, *Alactagulus pygmaeus* (lesser five-toed jerboa). Pedetidae: *Pca*, *Pedetes capensis* (springhare); Anomaluridae: *Asp*, *Anomalurus* sp. (scaly-tailed flying squirrel). Castoridae: *Cfi*, *Castor fiber* (Eurasian beaver). Geomyidae: *Tbo, Thomomys bottae* (pocket gopher). Caviidae: *Cpo*, *Cavia porcellus* (guinea pig). Hydrochoeridae: *Hhy*, *Hydrochoerus hydrochaeris* (capybara). Dasyproctidae: *Mac*, *Myoprocta acouchy* (acouchi). Thryonomyidae: *Tgr*, *Thryonomys gregorianus* (lesser cane rat). Octodontidae: *Ode*, *Octodon degus* (degu). Myocastoridae: *Mco*, *Myocastor coypus* (nutria). Hystricidae: *Hin*, *Hystrix indica* (Indian crested porcupine). Gliridae: *Dni*, *Dryomys nitedula* (forest dormouse). Sciuridae: *Mca*, *Marmota caudate* (long-tailed marmot); *Sfu*, *Spermophilus fulvus* (yellow ground squirrel); *Sca*, *Sciurus carolinensis* (gray squirrel); *Tas*, *Tamias asiaticus* (Asian chipmunk); and *Mbe*, *Menetes berdmorei* (Indochinese ground squirrel). Aplodontidae: *Aru*, *Aplodontia rufa* (mountain beaver); Primates: *Hsa*, *Homo sapiens* (human). Chinchillidae: *Cla*, *Chinchilla laniger* (long-tailed chinchilla). Lagomorpha: *Ocu*, *Oryctolagus cuniculus* (rabbit). The first two dots contain *M. musculus* DNA (250 and 500 ng, respectively); for the rest species, DNA samples amount to 500 ng (except for chinchilla, 150 ng). Hybridization of human DNA is due to *Alu*. Filters were hybridized and washed under mild conditions (60 and 42°C, respectively) for detecting both insignificantly and considerably divergent SINE copies.

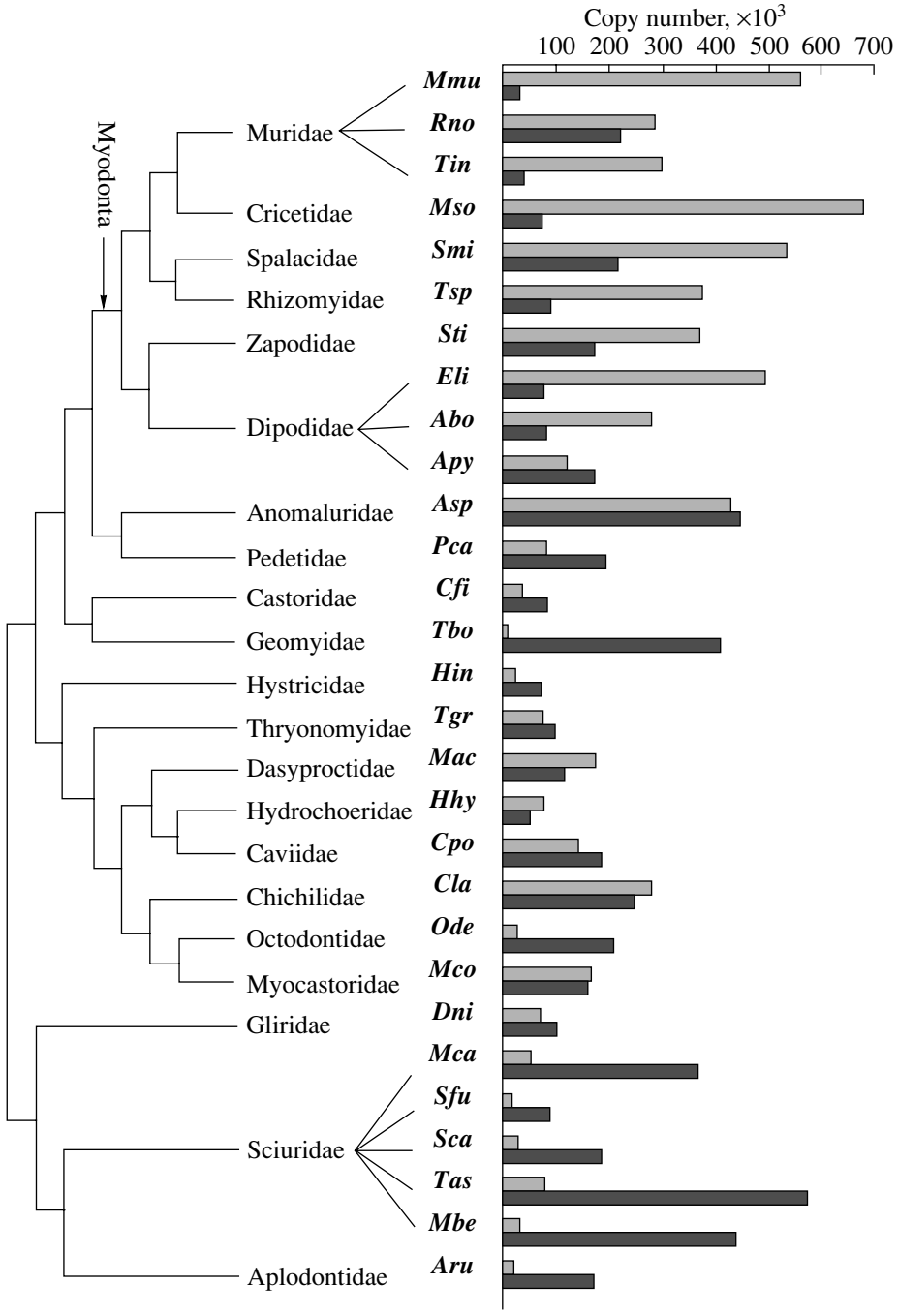

Fig. 2. Copy numbers of ID (black columns) and B1 (gray columns) in rodent genomes according to the dot hybridization data. The tree demonstrating the phylogenetic relationships of the rodent families under study was constructed using the data of several molecular phylogenetic works [9, 18, 21–25].

and, third, a joining of these elements into new SINEs is a widespread mechanism of their evolution.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research (projects nos. 05-04-49553 and 07-04-00462).

REFERENCES

1. Kramerov D., Vassetzky N. 2005. Short retroposons in eukaryotic genomes. *Int. Rev. Cytol*. **247**, 165–221.
2. Kramerov D.A., Grigoryan A.A., Ryskov A.P., et al. 1979. Long double-stranded sequences (dsRNA-B) of nuclear pre-mRNA consist of a few highly abundant classes of sequences: Evidence from DNA cloning experiments. *Nucleic Acids Res*. **6**, 697–713.

3. Krayev A.S., Kramerov D.A., Skryabin K.G., et al. 1980. The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA. *Nucleic Acids Res*. **8**, 1201–1215.
4. Deininger P.L., Jolly D.J., Rubin C.M., et al. 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J. Mol. Biol*. **151**, 17–33.
5. Haynes S.R., Toomey T.P., Leinwand L., et al. 1981. The Chinese hamster Alu-equivalent sequence: A conserved highly repetitious, interspersed deoxyribonucleic acid sequence in mammals has a structure suggestive of a transposable element. *Mol. Cell. Biol*. **1**, 573–583.
6. Ullu E., Tschudi C. 1984. *Alu* sequences are processed 7SL RNA genes. *Nature*. **312**, 171–172.
7. Quentin Y. 1994. Emergence of master sequences in families of retroposons derived from 7sl RNA. *Genetica*. **93**, 203–215.
8. Vassetzky N.S., Ten O.A., Kramerov D.A. 2003. B1 and related SINEs in mammalian genomes. *Gene*. **319**, 149–160.
9. Veniaminova N.A., Vassetzky N.S., Lavrenchenko L.A., et al. 2007. Reconstructing the phylogeny of rodents (Rodentia) from the results of structural analysis of short reproposon B1. *Genetika*. **43**, 916–929.
10. Kass D.H., Kim J., Deininger P.L. 1996. Sporadic amplification of ID elements in rodents. *J. Mol. Evol*. **42**, 7–14.
11. 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.
12. Martignetti J.A., Brosius J. 1993. Neural BC1 RNA as an evolutionary marker: Guinea pig remains a rodent. *Proc. Natl. Acad. Sci. USA*. **90**, 9698–9702.
13. Wang H., Iacoangeli A., Lin D., et al. 2005. Dendritic BC1 RNA in translational control mechanisms. *J. Cell. Biol*. **171**, 811–821.
14. Borodulina O.R., Kramerov D.A. 2005. PCR-based approach to SINE isolation: Simple and complex SINEs. *Gene*. **349**, 197–205.
15. Waterston R.H., Lindblad-Toh K., Birney E., et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*. **420**, 520–562.
16. Gibbs R.A., Weinstock G.M., Metzker M.L., et al. 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*. **428**, 493–521.
17. Serdobova I.M., Kramerov D.A. 1998. Short retroposons of the B2 superfamily: Evolution and application for the study of rodent phylogeny. *J. Mol. Evol*. **46**, 202–214.
18. Kramerov D.A. 1999. Evidence for a close phylogenetic relationship between the families Gliridae and Sciuridae based on the study of short reproposon B1-dID. *Dokl. Akad. Nauk*. **364**, 277–280.
19. Vasetskii N.S., Gogolevskaia I.K., Borodulina O.R., Kramerova D.A. 1999. B1-1D—A new Short retroposon from redeuts. *Mol. Biol*. (Moscow). **33**, 520–527.
20. Gogolevsky K.P., Kramerov D.A. 2006. Short interspersed elements (SINEs) of the Geomyoidea superfamily rodents. *Gene*. **373**, 67–74.
21. Adkins R.M., Gelke E.L., Rowe D., et al. 2001. Molecular phylogeny and divergence time estimates for major rodent groups: Evidence from multiple genes. *Mol. Biol. Evol*. **18**, 777–791.
22. Huchon D., Douzery E.J. 2001. From the Old World to the New World: A molecular chronicle of the phylogeny and biogeography of hystricognath rodents. *Mol. Phylogenet. Evol*. **20**, 238–251.
23. Huchon D., Madsen O., Sibbald M.J., et al. 2002. Rodent phylogeny and a timescale for the evolution of Glires: Evidence from an extensive taxon sampling using three nuclear genes. *Mol. Biol. Evol*. **19**, 1053–1065.
24. Adkins R.M., Walton A.H., Honeycutt R.L. 2003. Higher-level systematics of rodents and divergence time estimates based on two congruent nuclear genes. *Mol. Phylogenet. Evol*. **26**, 409–420.
25. Steppan S., Adkins R., Anderson J. 2004. Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst. Biol*. **53**, 533–553.