ISSN 0026-8933, Molecular Biology, 2006, Vol. 40, No. 1, pp. 63–71. © Pleiades Publishing, Inc., 2006.
Original Russian Text © O.A. Ten, O.R. Borodulina, N.S. Vassetzky, N.Iu. Oparina, D.A. Kramerov, 2006, published in Molekulyarnaya Biologiya, 2006, Vol. 40, No. 1, pp. 74–83.

GENOMICS. TRANSCRIPTOMICS. PROTEOMICS

UDC 577.21

Repetitive Sequences of the Tree Shrew Genome (Mammalia, Scandentia)

O. A. Ten, O. R. Borodulina, N. S. Vassetzky, N. Iu. Oparina, and D. A. Kramerov

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

Received July 21, 2005

Abstract—Copies of two repetitive elements of the common tree shrew (*Tupaia glis*) genome were cloned and sequenced. The first element, Tu III, is a ~260 bp long short interspersed element (SINE) with the 5' end derived from glycine RNA. Tu III carries long polypurine- and polypyrimidine-rich tracts, which may contribute to the specific secondary structure of Tu III RNA. This SINE was also found in the genome of the smooth-tailed tree shrew of another genus (*Dendrogale*). Tu III appears to be confined to the order Scandentia since it was not found in the DNA of other tested mammals. The second element, Tu-SAT1, is a tandem repeat with a monomer length of 365 bp. Some properties of its nucleotide sequence suggest that Tu-SAT1 is a centromeric satellite.

DOI: 10.1134/S0026893306010109

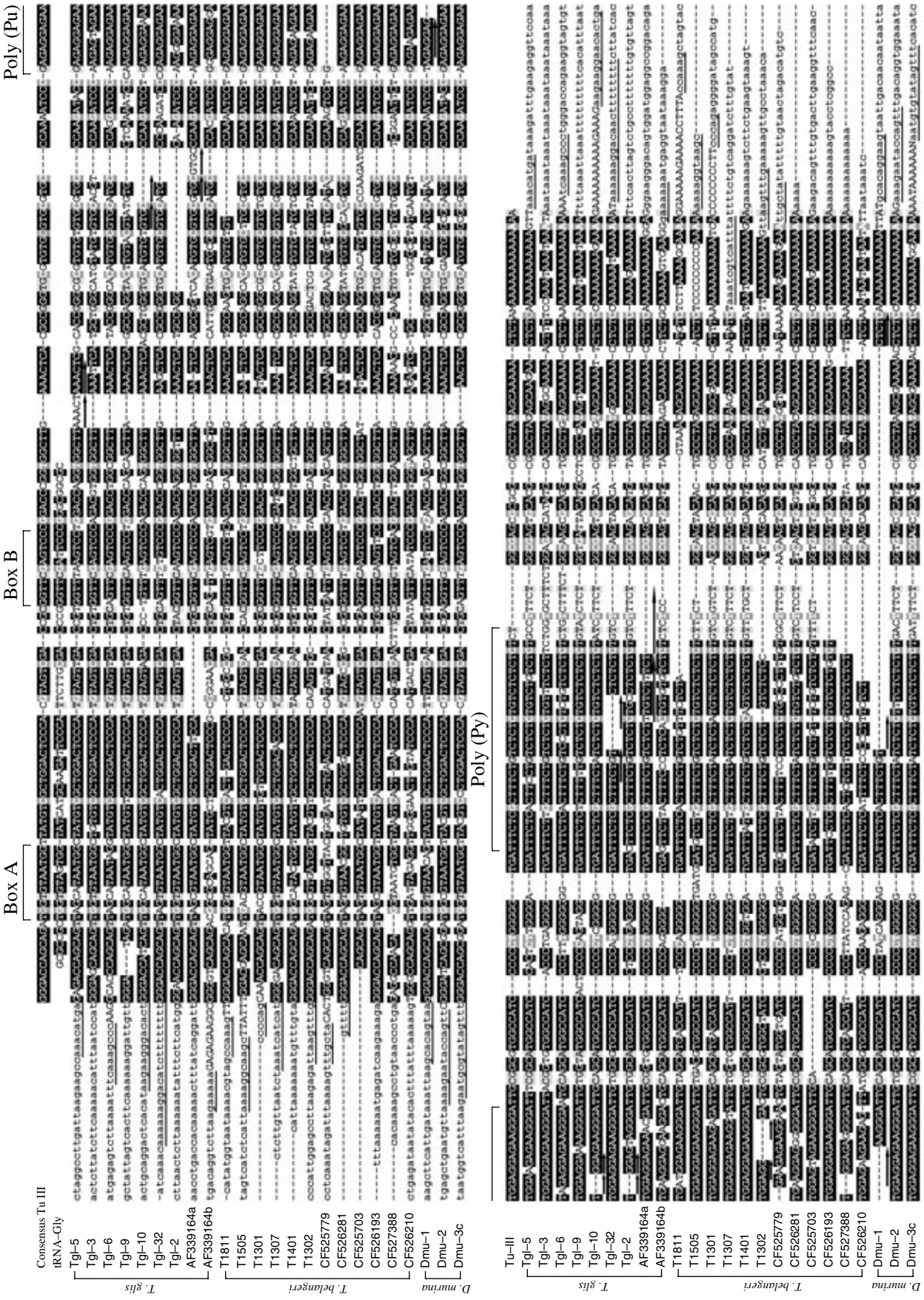
Key words: tree shrew, Scandentia, retrotransposon, SINE, tandem repeats, satellite DNA, secondary RNA structure

INTRODUCTION

Repetitive sequences make up a considerable fraction of eukaryotic genomes and can be divided into two main classes, tandem and interspersed repeats. Tandem repeats (satellite DNA) occupy extended regions in the centromeric and pericentromeric regions of chromosomes. The length of the repetitive unit (monomer) of satellite DNA can vary from 5 to over 200 bp in different species [1, 2]. The number of satellite monomers in the genome is commonly counted in millions. Monomer copies are much more similar within a species than between related species, which is due to concerted evolution based on unequal crossing over [3, 4]. Satellite DNAs rapidly replace one another so that species from different genera (e.g., *Rattus* and *Mus*) sometimes have no common satellite sequences [5, 6]. Although the significance of satellite DNAs is not quite clear, they seem to underlie the formation of chromosomal centromeres, and thus are probably involved in mitosis and meiosis [7, 8].

Interspersed repetitive sequences of the genome are largely represented by various mobile genetic elements including DNA transposons, retrotransposons (or LTR transposons), long interspersed elements (LINEs), and short interspersed elements (SINEs) [9, 10]. Except for DNA transposons, mobile elements use reverse transcription to amplify in the genome. Retrotransposons and LINEs use their own reverse transcriptases, while SINEs use the reverse transcriptase of partner LINEs. Mammalian genomes have a single family of active LINEs, L1 (apart from bovine Bov-B), whereas SINEs are variable, and over 20 active SINE families are presently known. SINEs are much shorter than other full-size mobile genetic elements: 100–400 bp vs. 4–7 kb. Commonly, the genome harbors one to three SINE families [11] and each of them is commonly represented by 10^5–10^6 copies, which is much greater than the number of other mobile genetic elements. Except for Alu from primates and B1 from rodents, which originated from 7SL RNA (small cytoplasmic RNA involved in the synthesis of secreted proteins), and SINE3 from zebrafish, which originated from 5S rRNA, SINEs are derived from various tRNAs. The 5' end of such SINEs is similar to tRNA. This region is followed by a region of unknown origin, specific for each SINE family. The 3' end of SINEs carries an A-rich region (for mammals and plants) or a region of several short direct repeats (for reptiles and fishes). Many properties of SINEs and other mobile genetic elements point to their selfish (parasitic) nature [12]. However, the existence of SINE families in the genome for many millions of years has made them an important factor in evolution. For instance, their integration in the vicinity of genes or into their introns can alter the transcription and splicing of genes [13]. SINEs can also play an important role in the protection of cells from stress [14].

SINEs have proven to be convenient tools for phylogenetic analysis. Both SINE families [15] and individual SINE copies integrated into particular loci [16, 17]


Box A
Box B
Poly (Pu)
Poly (Py)
Consensus Tu III
tRNA–Gly
Tgl–5
Tgl–3
Tgl–6
Tgl–9
Tgl–10
Tgl–32
Tgl–2
AF339164a
AF339164b
T1811
T1505
T1301
T1307
T1401
T1302
CF525779
CF526281
CF525703
CF526193
CF527388
CF526210
Dmu–1
Dmu–2
Dmu–3c
Tu–III
T. glis
T. belangeri
D. murina

can be used as phylogenetic markers. This approach is substantiated by the fact that the emergence of the same SINE families or SINE integration into the same genomic locus are highly unlikely. SINEs have been used to study the evolution of humans [18], rodents [19], artiodactyls and cetaceans [16], the mammalian clades Afrotheria [20] and Euarchontoglires [21], etc.

In this study, the most abundant repetitive sequences from the genome of the tree shrew *Tupaia glis* were cloned and described. These were represented by a satellite DNA and a SINE family. Studies on the genome of tree shrews and their repetitive sequences are of apparent interest. Initially, tree shrews were assigned to the order Insectivora. Then they were considered to be a family of the order Primates (this viewpoint is still popular in the Russian scientific literature). Later, tree shrews were classified as a separate order Scandentia [22, 23]. The latter classification seems most justified. Further studies, in particular, on nuclear genes and repeated sequences, are required to definitely resolve the relationships between primates, tree shrews, flying lemurs, rodents, and lagomorphs, all of which form the clade Euarchontoglires. Exploration of the genome of tree shrews is also of interest considering their use as laboratory animals in physiological and medical studies. The convenience of tree shrews for this role is due to the fact that they are more similar to humans than are other laboratory animals, with the exception of primates (the keeping of which is more expensive and complicated).

EXPERIMENTAL

Ethanol-fixed tissues of *Tupaia glis* (Vietnam, Dalat environs) and *Dendrogale murina* were kindly provided by O. Likhnova (Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences) and V. Matveev (Moscow State University). DNA was isolated using proteinase K digestion and phenol–chloroform extraction.

Cloning and selection of repetitive sequences. Genomic DNA was hydrolyzed by *Hind*III and *Eco*RI and cloned into pGEM7Zf. Transformed *E. coli* XL1-Blue colonies were screened by hybridization with total labeled genomic DNA from *T. glis.* DNA was labeled with [α^{32}-P]dATP by random priming (Prime-ItII, Stratagene, United States). The hybridization conditions have been described elsewhere [24]. Twenty positive clones were selected and the inserted DNA fragments were analyzed by cross hybridization, which allowed all clones to be divided into two groups (containing SINE and satellite DNA, respectively; see below). The insertion fragments were sequenced either immediately, or after the subcloning of their restriction fragments containing repetitive sequences (as tested by blot hybridization).

Monomers of satellite DNA from *T. glis* were cloned after digestion with *Bgl*II into pSL1109, and positive clones were selected by hybridization with a labeled DNA insertion fragment of a second group clone.

Sequencing. Nucleotide sequences were determined by a modified dideoxynucleotide method with Sequenase 2.0 (Amersham, United States) and [α^{32}-P]dATP [25]. Standard M13 primers or tree shrew SINE-specific primers (for long cloned fragments; Tgl-SINE.dir: 5'-TGGYTATGTYGCTG-GACTCCCAT-3' and Tgl-SINE.rev: 5'-TGGCCATR-CYRCAAATCCCTT-3') were used. The sequences were deposited in GenBank (SINE, accession nos. AY661722–AY661731; satellite, AY661714–AY661721).

Polymerase chain reaction (PCR) was used to study the distribution of Tu III SINE in mammals. The reaction mixture contained 0.1 ng of genomic DNA from tree shrews (*T. glis* or *D. murina*) or 10 ng of DNA from other mammals. Tgl-SINE.dir and Tgl-SINE.rev were used as primers. Twenty five PCR cycles were performed, each consisting of 1 min at 95, 64, and 72°C. PCR products were separated by electrophoresis in 3% agarose gel (2% NuSieve and 1% agarose) and were stained with ethidium bromide. Amplified fragment of DNA from *D. murina* was eluted from the gel and cloned into pGEM-T (Promega, United States).

Blot hybridization of satellite DNA. Genomic DNA from *T. glis* was digested with *Bgl*II to produce both partially and completely digested DNA. After separation in 1% agarose gel, DNA was stained, transferred onto a nylon membrane by capillary blotting, and hybridized with a ^{32}P-labeled monomer of satellite DNA cloned from *T. glis.*

Computer analysis of sequences. In the consensus sequence of Tu III SINE, methylation-dependent hypervariable sites CpG/TpG/CpA were converted to CpG.

Fig. 1. Alignment of nucleotide sequences of Tu III SINE from the genomes of tree shrews *Tupaia glis, T. belangeri*, and *Dendrogale murina.* Tgl and Dmu copies were sequenced in this study, while other sequences were extracted from GenBank (genomic copies T1301, T1302, T1307, T1401, T1505, T1811 [29] as well as cDNA sequences CF^{nnn}). Tu III consensus was constructed from these copies, 36 EST sequences, and 12 sequences of PCR products from *D. murina* obtained in this study. The nucleotide sequences of Tu III are shown in uppercase, while target site duplications are shown in underlined lower case letters. Arrows underline the direct repeats within Tu III resulting from the duplications or lost after the deletions. The polypurine and polypyrimidine tracts are indicated above. Glycine tRNA^{CCC} from the database of genomic tRNA GtRNAdb (human chr16#34) [32] is shown as a possible evolutionary precursor of Tu III.

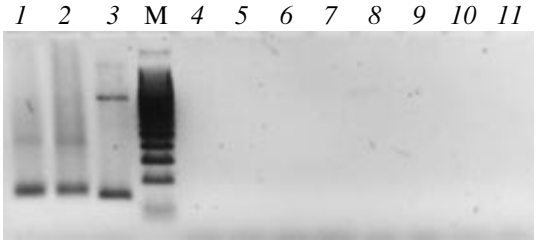


Fig. 2. Tu III distribution in mammalian genomes. PCR was conducted using Tu III-specific primers. The reaction mixture contained 0.1 ng of genomic DNA from the tree shrew or 10 ng of genomic DNA from other mammals: (*1*) common shrew (*Tupaia glis*), (*2*) northern smooth-tailed tree shrew (*Dendrogale murina*), (*3*) Tu III-containing clone (Tgl-10), (*4*) Daurian hedgehog (*Mesechinus dauricus*), (*5*) mouse (*Mus musculus*), (*6*) rabbit (*Oryctolagus cuniculus*), (*7*) dog (*Canis familiaris*), (*8*) Bennett's tree kangaroo (*Dendrolagus bennettianus*), (*9*) cow (*Bos taurus*), and (*10*) human (*Homo sapiens*); M, 100-bp DNA ladder.

The secondary structure of RNA was predicted using mfold 3.2 software [26] with standard parameters. DNA conformation was simulated using the Plot_it (http://hydra.icgeb.trieste.it/~kristian/dna/plot_it.html) and TwistFlex (http://margalit.huji.ac.il/TwistFlex/index.html) servers using a window of 30 bp. Changes in B-DNA twisting [27] and total DNA flexibility were analyzed with default parameters for trinucleotides [28].

RESULTS AND DISCUSSION

Previously, we searched the genome of common tree shrew *T. glis* for 7SL RNA-derived SINEs in order to confirm the relationship between rodents, primates, and tree shrews [21]. Indeed, such SINEs were found, although the number of their copies in the tree shrew genome (500–1000) was much lower than the number of B1 copies in the mouse genome (10^5) or Alu copies in the human genome (~10^6). Later we screened the genomic library of *T. glis* with the labeled genomic DNA of this species, rather than with the 7SL-specific probe. This yielded a number of clones that also hybridized with the second probe, which indicated the presence of highly repetitive nucleotide sequences among the cloned DNA fragments. In this study, these clones were divided into two groups by cross hybridization. DNA sequencing identified an interspersed repeat in the clones of one group and a tandem repeat in the clones of the other group.

Major SINE of Tree Shrew Genome

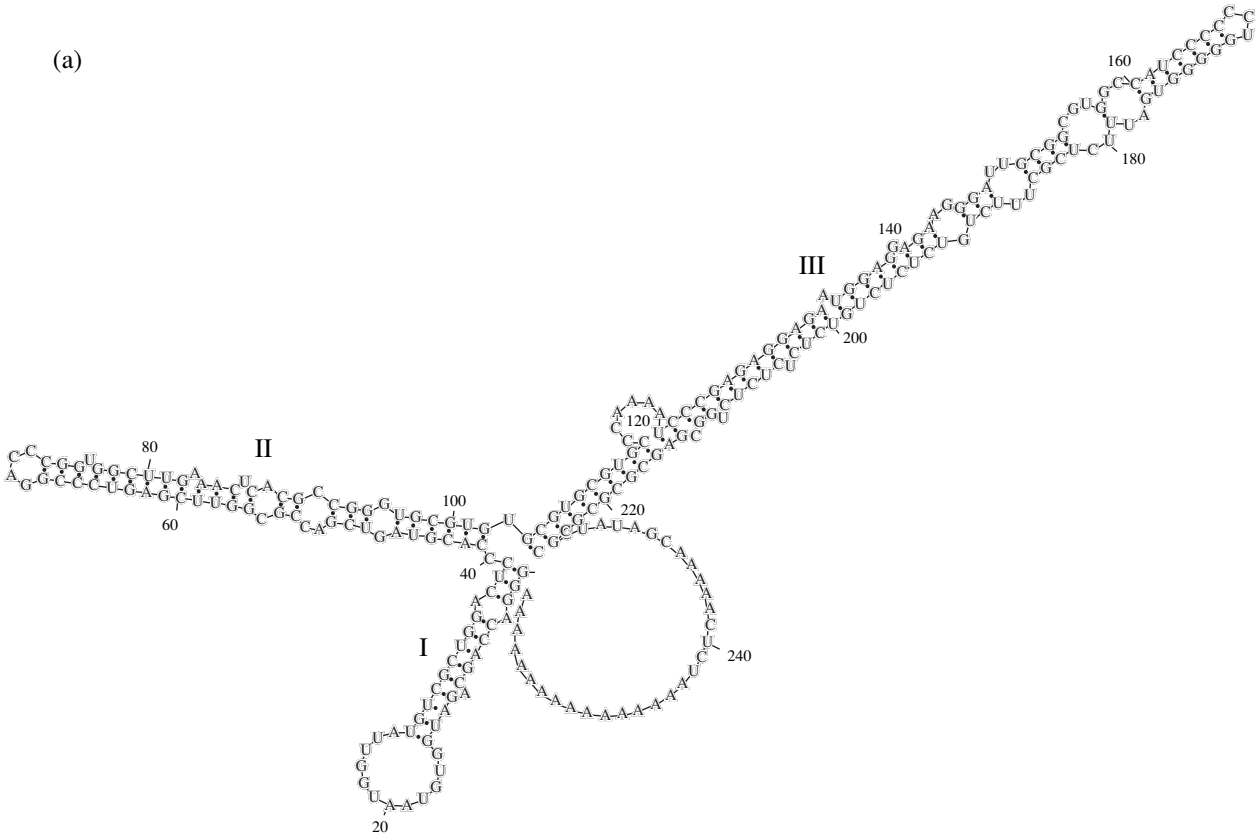
Figure 1 shows the sequence similarity for clones of the first group (seven upper sequences). The relative shortness of the similarity region (260 bp), the A and B boxes of RNA polymerase III promoter in the 5'-terminal region, the A-rich sequence in the 3'-terminal region, and the short flanking repeats indicate that the revealed repetitive sequence is a typical SINE. A search for similarity in nucleotide databases has revealed two unannotated copies of this SINE in an intron of the *M6P/IGF2R* gene of *T. glis* (AF339164) and six more copies (AB090261–66) in the genome of another tree shrew species—*T. belangeri*—described as Tu type III SINE [29]. This publication gives only a brief description of this SINE, while focusing on two low-copy-number (~200 copies per genome) SINEs Tu type I and II with monomers derived from both tRNA and 7SL RNA. In addition, 6 full-size and 36 truncated nucleotide sequences of Tu III SINE annotated as fragments of mRNA from the brain of *T. belangeri* were identified in the EST database. These full-size sequences were aligned with all the above-mentioned genomic copies of Tu III SINE and are shown in Fig. 1. The differences between Tu III copies of different species of the genus *Tupaia* did not exceed intraspecific differences.

The range of Tu III distribution was determined by PCR amplification of genomic DNA from certain mammals. PCR products of expected size (140 bp) were observed in tree shrews of both genera (*Tupaia* and *Dendrogale*) but not in other tested mammalian orders (Fig. 2). This indicates the presence of Tu III SINE in the order Scandentia and its absence from the genomes of other mammals. The PCR product from *D. murina* was cloned and sequenced from 12 clones. The resulting nucleotide sequences corresponded to the central region of Tu III SINE (data not shown). In order to unambiguously confirm the presence of a typical Tu III SINE in the genome of *D. murina*, the genomic library of this tree shrew species was screened with a Tu III-specific probe and three positive clones were sequenced. This allowed us to reveal Tu III copies of *D. murina* (Fig. 1) and to conclude that the Tu III SINEs from the genomes of two tree shrew species are very similar. In all likelihood, this SINE is typical for all tree shrews but its distribution is limited to the order Scandentia.

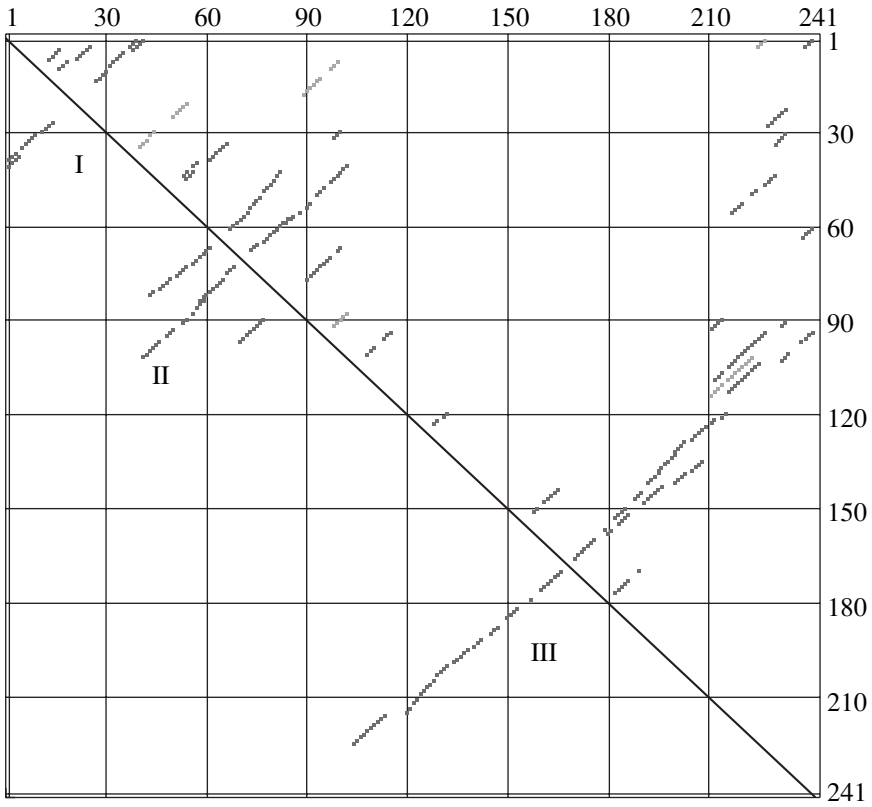
Nishihara *et al.* failed to identify the tRNA species which was the evolutionary precursor of Tu III [29]. Another approach allowed us to reveal a 70% similarity between the 5'-terminal region of Tu III and a vari-

Fig. 3. Secondary structure of Tu III RNA. (a) a variant of folding of the consensus sequence Tu III RNA with the least free energy $\Delta G(-105)$. Three hairpin structures are designated by Roman numerals. (b) analysis of conserved elements in predicted secondary structures. Segments of dots above the diagonal correspond to hairpin structures in individual variants of the predicted Tu III RNA structure, while similar segments below the diagonal represent hairpin structures conserved in all or most variants. All three hairpins in the above secondary structure are conserved, although the region of hairpin II can be folded in different ways.

(a)


I
II
III
20
40
60
80
100
120
140
160
180
200
220
240

(b)


1
30
60
90
120
150
180
210
241
I
II
III

ant of glycine $tRNA^{CCC}$ (Fig. 1). Accordingly, this tRNA can be considered as a possible evolutionary precursor of Tu III.

The tRNA-unrelated region of Tu III starts with a 42 bp region without specific features except the presence of a $(YGTG)_3$ sequence. This region is followed by a 24-bp polypurine tract and, 31 bp downstream of it, a 38-bp polypyrimidine tract with only three purines. Similar to other mammalian SINEs, Tu III ends with an A-rich tail. Most copies are flanked with short direct repeats (Fig. 1) resulting from target site duplication during the integration.

While polypyrimidine tracts can be found in many mammalian SINEs [24], no internal (non-tail) polypurine tracts have been reported. Both these tracts in Tu III included 2–3 copies of short tandem repeats; the number of these repeats could sometimes vary (underlined in Fig. 1). Such variability could be due to slippage of DNA polymerase during replication. Such a mechanism of deletions and insertions is well known for micro- and minisatellite DNA [30]. Recent analysis of the structure and evolution of B1 SINE allowed us to reveal internal duplications and deletions flanked by very short (3–6 bp) direct repeats [21]. We proposed that such repeats promote duplications and deletions through the jumping of reverse transcriptase from one repeat to another during the synthesis of SINE RNA from a DNA template [11]. Analysis of Tu III sequences also demonstrated several deletions whose formation could be mediated by short repeats located at a certain distance from each other (Fig. 1, Tgl_32 and Dmi_1, underlined). Note that one of the repeats is always within a deletion. It is not improbable that the above-mentioned tandem repeats of Tu III also arose or disappeared during reverse transcription.

Figure 3a shows a variant of the predicted secondary structure of Tu III RNA with the least free energy ($\Delta G = -105$). The tRNA-related region of Tu III forms two hairpins, while the central region is folded into a long hairpin with just a few short unpaired regions. Comparison of different predicted secondary structures of Tu III RNA with free energy close to the minimum (Fig. 3b) demonstrated the presence in all of them of this long hairpin, largely resulting from pairing between the polypurine and polypyrimidine tracts. This secondary structure may be important for the functioning of Tu III RNA.

Major Satellite DNA from the Genome of *Tupaia glis*

The genomic library produced by ligation of *Hin*dIII and *Eco*RI digested genomic DNA from *T. glis* into a plasmid vector allowed a considerable number of clones with a sequence different from Tu III SINE to be isolated. Although the insert size considerably varied between these clones, their partial sequencing demonstrated the presence of very similar and apparently tandem DNA repeats (data not shown). Since the cloned fragments were too large for complete sequencing using standard primers alone, we generated a genomic library from *T. glis* DNA digested with *Bgl*II. This enzyme was selected because each monomer of the revealed tandem repeat included its recognition site. After hybridization screening of the library, the DNA insertion fragments of seven clones were sequenced (Fig. 4). Six of them had similar sizes (361–367 bp) and nucleotide sequences. The highest sequence similarity was observed between clones Tu_SAT-07, -13, -14, and -17 (group 1) as well as between Tu_SAT-06 and -11 (group 2). The fragment of clone Tu-SAT-25 differed more, had a disturbed *Bgl*II site, and was accordingly longer. An "additional" region of this fragment (Tu_SAT-25') also aligned with other cloned sequences (Fig. 4), which supported a tandem organization of this repeat.

This was further confirmed by complete and partial hydrolysis of genomic DNA from *T. glis* by *Bgl*II and subsequent agarose gel electrophoresis (Fig. 5a) and blot hybridization (Fig. 5b). An observed ladder of hybridized fragments is typical of tandem repeats. Hence, the DNA sequence revealed in the genome of common tree shrew appears to be a satellite DNA.

Analysis of this DNA designated as Tu-SAT1 demonstrated the absence of similarity with known satellite DNA from other species, which is typical for centromeric satellites. Despite a relatively high A+T content (61.7%) typical for satellite DNA, the nucleotide sequences of Tu-SAT1 included GC-rich segments (GGCCC) (Fig. 4). At least some of these regions can interact with proteins. For instance, the 17-bp motif recognized by centromeric protein CENP-B is the GC-richest fragment of primate α satellite [31].

The nucleotide sequence of the Tu-SAT1 satellite is saturated with short simple repeats with mono- and dinucleotide units, which is also typical of satellite DNA (Fig. 4). At least some satellites were proposed to be formed by duplications of a short repeat. In this case, degenerate variants of it can be found along the whole satellite monomer (e.g., the 5 bp repeat in α satellite DNA of primates). Despite the presence of single direct and inverted repeats (Fig. 4), such an origin of Tu-SAT1 could not be confirmed.

Analysis of the Tu-SAT1 sequence demonstrated a high level of DNA flexibility and the presence of alternating segments with different flexibility in the monomer. Such an irregular structure is common among complex satellite DNA including among centromeric satellites.

It is also important to note that the length of Tu-SAT1 (360–367 bp) roughly corresponds to the length of two nucleosomes, which is also typical of satellite

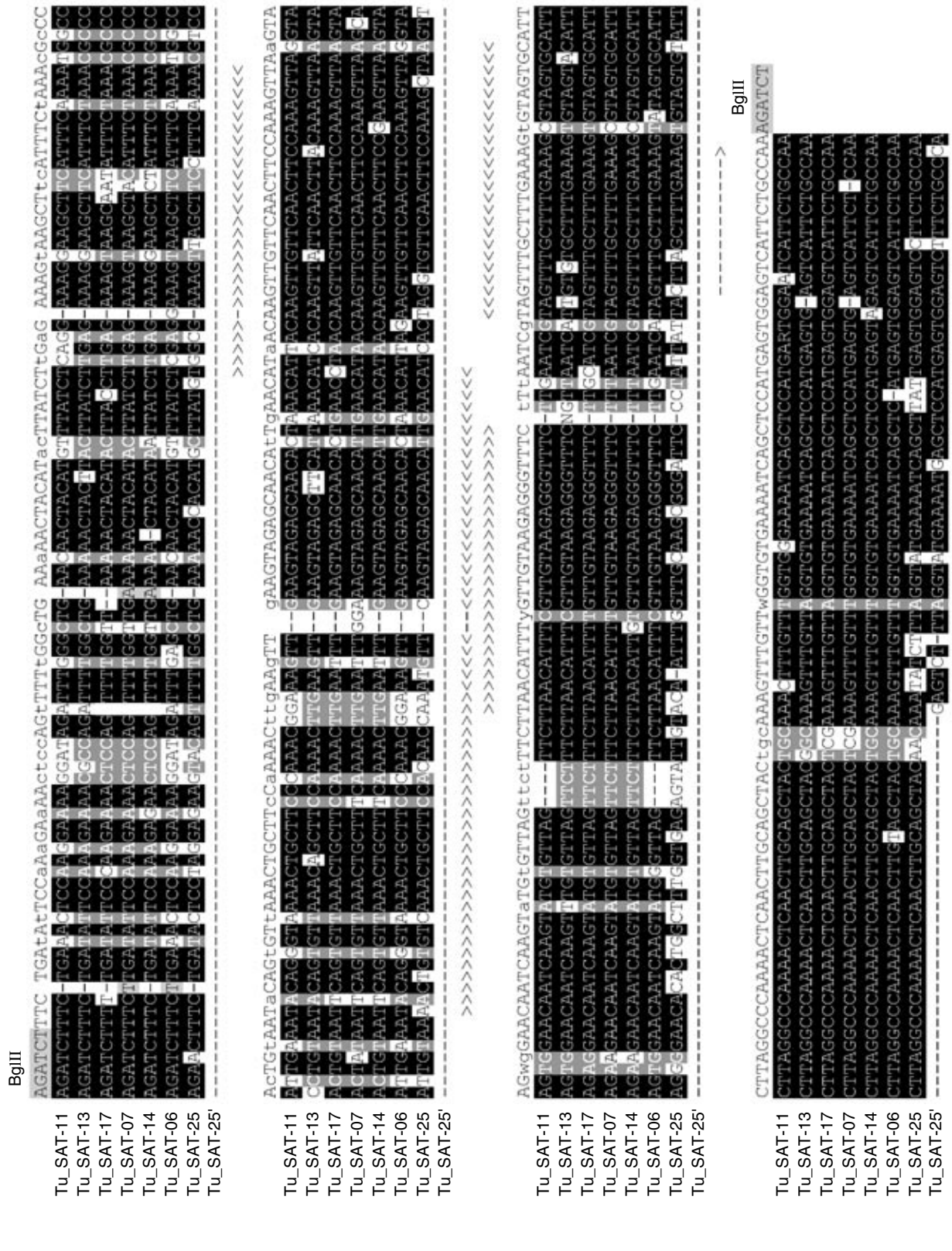


Fig. 4. Alignment of *Bgl*II monomers of Tu-SAT1 repeat from the genome of the common tree shrew *Tupaia glis*. The nucleotide sequences from seven clone monomers and their consensus sequence are shown. The right *Bgl*II site is not shown. The left *Bgl*II site of the Tu_SAT-25 clone was disturbed by a mutation and the cloned sequence is accordingly longer. The additional sequence (Tu_SAT-25′) can be aligned with other monomers, which confirms a tandem organization of Tu-SAT1 repeats. Inverted and direct repeats are designated as >>>/<<< and – – –>, respectively.

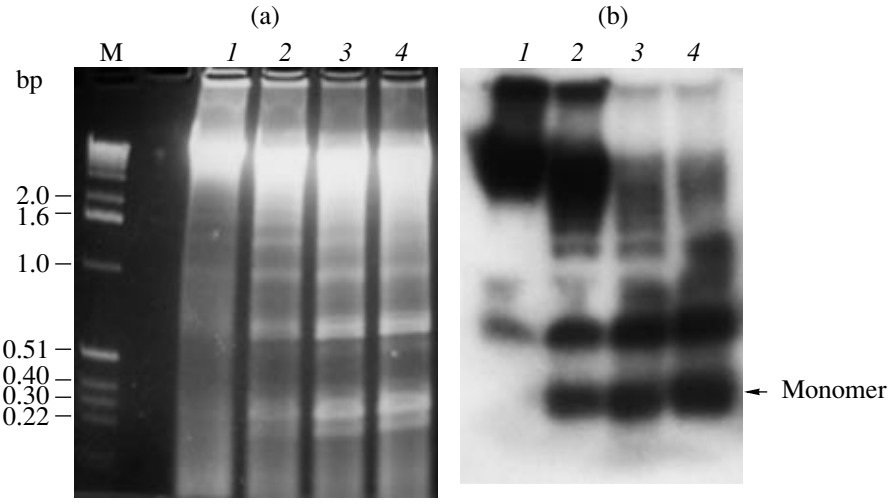


Fig. 5. Tandem organization of the Tu-SAT1 repeat. Genomic DNA of *Tupaia glis* was digested with *Bgl*II for (*1*) 5, (*2*) 15, (*3*) 30, and (*4*) 60 min; (a) DNA electrophoresis in 1.5% agarose gel; (b) hybridization of DNA transferred onto a nylon membrane with a labeled Tu-SAT1-specific probe. The observed ladder confirms the tandem organization of Tu-SAT1 monomers. M, DNA size marker (Gibco BRL, United States).

DNA (171 bp in human, 180 bp in rice, 340 bp in beet, etc.) [5, 6].

In this work, two major genomic repeats of the tree shrew were described. One of them, Tu III, is a typical SINE, although it has some specific features—in particular, it contains a polypyrimidine tract and the transcript of this SINE has a specific secondary structure. The distribution of this repeat is limited to tree shrews (Scandentia), which confirms the monophyly of this order. The second repeat, Tu-SAT1, demonstrates the properties of a typical centromeric satellite. The finding of repetitive sequences can serve as a starting point for studying the genome of this promising laboratory animal.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 05-04-49553 and 02-04-48644).

REFERENCES

1. Beridze T.G. 1982. *Satellitnaya DNK* (Sattelite DNA). Moscow: Nauka.
2. Willard H.F. 1990. Centromeres of mammalian chromosomes. *Trends Genet*. **6**, 410–416.
3. Elder J.F., Jr., Turner B.J. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. *Quant. Rev. Biol*. **70**, 297–320.
4. Liao D. 1999. Concerted evolution: Molecular mechanism and biological implications. *Am. J. Hum. Genet*. **64**, 24–30.
5. Henikoff S., Ahmad K., Malik H.S. 2001. The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science*. **293**, 1098–1102.
6. Ugarkovic D., Plohl M. 2002. Variation in satellite DNA profiles: Causes and effects. *EMBO J*. **21**, 5955–5959.
7. Lamb J.C., Birchler J.A. 2003. The role of DNA sequence in centromere formation. *Genome Biol*. **4**, 214.
8. Enukashvili N.I., Kuznetsova I.S., Podgornaya O.I. 2003. The mammalian centromere organization. *Tsitologiya*. **45**, 255–270.
9. Smit A.F. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Current Opin. Genet. Dev*. **9**, 657–663.
10. Deininger P.L., Batzer M.A. 2002. Mammalian retroelements. *Genome Res*. **12**, 1455–1465.
11. Kramerov D.A., Vassetzky N.S. 2005. Short retroposons (SINEs) in eukaryotic genomes. *Int. Rev. Cytol*. **247**, 165–221.
12. Orgel L.E., Crick F.H. 1980. Selfish DNA: The ultimate parasite. *Nature*. **284**, 604–607.
13. Makalowski W. 2000. Genomic scrap yard: How genomes utilize all that junk. *Gene*. **259**, 61–67.
14. Allen T.A., von Kaenel S., Goodrich J.A., Kugel J.F. 2004. The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nature Struct. Mol. Biol*. **11**, 816–821.
15. Serdobova I.M., Kramerov D.A. 1994. Use of short retroposons as phylogenetic markers. *Dokl. Akad. Nauk*. **335**, 664–667.
16. Shimamura M., Yasue H., Ohshima K., Abe H., Kato H., Kishiro T., Goto M., Munechika I., Okada N. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. *Nature*. **388**, 666–670.
17. Shedlock A.M., Okada N. 2000. SINE insertions: Powerful tools for molecular systematics. *Bioessays*. **22**, 148–160.
18. Stoneking M., Fontius J.J., Clifford S.L., Soodyall H., Arcot S.S., Saha N., Jenkins T., Tahir M.A., Deininger P.L., Batzer M.A. 1997. Alu insertion polymorphisms and

human evolution: evidence for a larger population size in Africa. *Genome Res*. **7**, 1061–1071.

19. Kramerov D., Vassetzky N., Serdobova I. 1999. The evolutionary position of dormice (Gliridae) in Rodentia determined by a novel short retroposon. *Mol. Biol. Evol*. **16**, 715–717.
20. Nikaido M., Nishihara H., Hukumoto Y., Okada N. 2003. Ancient SINEs from African endemic mammals. *Mol. Biol. Evol*. **20**, 522–527.
21. Vassetzky N.S., Ten O.A., Kramerov D.A. 2003. B1 and related SINEs in mammalian genomes. *Gene*. **319**, 149–160.
22. Carrol R.L. 1988. *Vertebrate Paleontology and Evolution*. N.Y.: Freeman and Company.
23. Pavlinov I.Ya. 2003. *Sistematika sovremennykh mlekopitayushchikh* (Taxonomy of Recent Mammals), Moscow: Mosk. Gos. Univ.
24. Borodulina O.R., Kramerov D.A. 2001. Short interspersed elements (SINEs) from insectivores. Two classes of mammalian SINEs distinguished by A-rich tail structure. *Mammal. Genome*. **12**, 779–786.
25. Redston M.S., Kern S.E. 1994. Klenow co-sequencing: A method for eliminating "stops." *Biotechniques*. **17**, 286–288.
26. Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. **31**, 3406–3415.
27. Sarai A., Mazur J., Nussinov R., Jernigan R.L. 1989. Sequence dependence of DNA conformational flexibility. *Biochemistry*. **28**, 7842–7849.
28. Gromiha M. 2000. Structure based sequence dependent stiffness scale for trinucleotides: A direct method. *J. Biol. Phys*. **26**, 43–50.
29. Nishihara H., Terai Y., Okada N. 2002. Characterization of novel Alu- and tRNA-related SINEs from the tree shrew and evolutionary implications of their origins. *Mol. Biol. Evol*. **19**, 1964–1972.
30. Debrauwere H., Gendrel C.G., Lechat S., Dutreix M. 1997. Differences and similarities between various tandem repeat sequences: Minisatellites and microsatellites. *Biochimie*. **79**, 577–586.
31. Masumoto H., Nakano M., Ohzeki J. 2004. The role of CENP-B and alpha-satellite DNA: *De novo* assembly and epigenetic maintenance of human centromeres. *Chromosome Res*. **12**, 543–556.
32. Schattner P., Brooks A.N., Lowe T.M. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res*. **33**, W686–W689.