A Novel Abundant Family of Retroposed Elements (DAS-SINEs) in the Nine-Banded Armadillo (*Dasypus novemcinctus*)

Gennady Churakov, Arian F.A. Smit,† Jürgen Brosius,* and Jürgen Schmitz**

*Institute of Experimental Pathology (ZMBE), University of Münster, Münster, Germany; and †Institute for Systems Biology, Seattle

About half of the mammalian genome is composed of retroposons. Long interspersed elements (LINEs) and short interspersed elements (SINEs) are the most abundant repetitive elements and account for about 21% and 13% of the human genome, respectively. SINEs have been detected in all major mammalian lineages, except for the South American order Xenarthra, also termed Edentata (armadillos, anteaters, and sloths). Investigating this order, we discovered a novel high-copy-number family of tRNA derived SINEs in the nine-banded armadillo *Dasypus novemcinctus,* a species that successfully crossed the Central American land bridge to North America in the Pliocene. A specific computer algorithm was developed, and we detected and extracted 687 specific SINEs from databases. Termed DAS-SINEs, we further divided them into six distinct subfamilies. We extracted $tRNA^{Ala}$-derived monomers, two types of dimers, and three subfamilies of chimeric fusion products of a $tRNA^{Ala}$ domain and an approximately 180-nt sequence of thus far unidentified origin. Comparisons of secondary structures of the DAS-SINEs' tRNA domains suggest selective pressure to maintain a tRNA-like D-arm structure in the respective founder RNAs, as shown by compensatory mutations. By analysis of subfamily-specific genetic variability, comparison of the proportion of direct repeats, and analysis of self-integrations as well as key events of dimerization and deletions or insertions, we were able to delineate the evolutionary history of the DAS-SINE subfamilies.

Introduction

Short interspersed elements (SINEs) are prominent components of many eukaryotic genomes. As a subclass of retronuon (a nuon is any distinct nucleic acid module), they show great exaptive potential as novel parts of genes, including regulatory elements (Brosius and Gould 1992). Humans harbor more than one million SINE elements covering 13% of the genome (Lander et al. 2001). SINEs are typically derived from tRNAs, with the exception of the primate Alu, the rodent B1, the Tupaia Tu type I, and II SINEs that originated from 7SL RNA (Ullu and Tschudi 1984; Rogers 1985; Szemraj et al. 1995; Nishihara, Terai, and Okada 2002), as well as 5S rRNA-derived elements (Kapitonov and Jurka 2003). tRNA-derived SINEs are abundant modules in a broad range of multicellular organisms, including vertebrates, invertebrates, and plants (Okada and Ohshima 1995). In mammals, tRNA-derived SINEs are described in the superordinal clade I (Afrotheria [Nikaido et al. 2003]), clade III (Rodentia [Kramerov et al. 1979], Lagomorpha: [Cheng et al. 1984], Primates [Daniels and Deininger 1991], Dermoptera [Schmitz and Zischler 2003], and Scandentia [Nishihara, Terai, and Okada 2002]), and clade IV (Laurasiatheria [Shimamura et al. 1997; Nikaido, Rooney, and Okada 1999; Bentolila et al. 1999; Borodulina and Kramerov 2001]). So far, none have been described in clade II (Xenarthra) (for classification see Murphy et al. [2001]). Here, we report the discovery of the first SINE family in xenarthrans.

The initial steps of retroposition require an RNA template. Reverse transcription is probably performed by a long interspersed element (LINE)–encoded endonuclease/reverse transcriptase, the only known eukaryotic source that performs reverse transcription in trans (Kajikawa and Okada 2002; Dewannieux, Esnault, and Heidmann 2003). Processed as well as unprocessed RNAs can serve as templates for reverse transcription (Schmitz et al. 2004). It has been shown that the genomic integration sites in the most abundant primate and rodent SINEs exhibit a preference for TTAAAA motifs (Jurka 1997). In addition, Schmitz et al. (2004) showed that the 3′ end of template RNAs can choose appropriate genomic targets by base complementarity.

tRNA-derived SINEs are found in different combinations. Monomeric tRNA-SINEs are present, for example, in the primate infraorder lorisiformes (Roos, Schmitz, and Zischler 2004), first discovered in *Galago crassicaudatus* (Galago monomer [Daniels and Deininger 1991]), in the dermopteran species *Cynocephalus variegatus* (CYN-I-SINE [Schmitz and Zischler 2003]), in rodents (ID-SINEs [Kass, Kim, and Deininger 1996]), and in the camelid species *Vicugna vicugna* (vic-1 [Lin et al. 2001]). Dimeric and trimeric tRNA-SINEs with identical subunits could be characterized in dermopterans (CYN-II, CYN-III [Schmitz and Zischler 2003], or t-SINE [Piskurek et al. 2003]). In many instances, tRNA-derived SINEs display fusion products of a tRNA-derived and a tRNA-unrelated subunit.

There are at least three conceivable mechanisms that generate combined tRNA-derived SINEs: (1) tRNA genes are frequently arranged in clusters and may, for instance, cotranscribe as dicistrons. One example of a retroposed, probably unprocessed RNA Pol III transcript, is the *Twin* SINE of the mosquito *Culex pipiens*. The unprocessed dicistronic transcript has been retroposed and serves as a master locus for new SINEs (Feschotte et al. 2001). Subunits of the *Twin* SINEs are separated by a 39-nt spacer that is assumed to be also present at the locus of the founder tRNA genes. (2) Genomic integration of a reverse transcribed monomeric element into the oligo(A) tail of a preexisting master SINE can lead to subsequent transcription of the new dimeric structure. Beside tRNA-derived SINEs, the most familiar example of dimerization is the fusion of 7SL-derived FLAM (free left Alu monomers) and FRAM (free right Alu monomers) to form Alu dimers (Quentin 1992). (3) Template switching during reverse

Key words: tRNA, SINE, armadillo, retroposition, SINE evolution.

E-mail: jueschm@uni-muenster.de.

Mol. Biol. Evol. 22(4):886–893. 2005
doi:10.1093/molbev/msi071
Advance Access publication December 22, 2004

transcription generates chimeric retronuons (Brosius 1999; Gilbert and Labuda 2000; Cost et al. 2002; Buzdin 2004).

Borodulina and Kramerov (2001) distinguished two structural variants of mammalian SINEs by comparing tRNA-related SINEs. Like all 7SL derivates and, for example, the tRNA-derived SINEs of dermopterans (CYN-SINEs), class T^- SINEs show a more or less homogeneous 3′ terminal oligo(A) segment. On the other hand, class T^+ SINEs, like most other tRNA-related SINEs, represent a more complex A-rich tail, including an AATAAA-TCTTT/$(T)_3A(_n)$-like motif. The conserved AATAAA motif has been suggested to act as a polyadenylation signal followed by an RNA Pol III termination signal (Borodulina and Kramerov 2001). Class T^- and class T^+ SINEs are discussed as possible outcomes of different mechanisms of retroposition (Borodulina and Kramerov 2001). Based on our findings from surveying the sequences of the armadillo genome we discuss the involvement of both mechanisms to account for different DAS subfamilies.

Methods

Strategy to Find Novel SINE Elements

To extract unknown transposed elements from GenBank entries, we queried genomic sequences from all mammalian genome-sequencing projects. For the nine-banded armadillo, we retrieved about 4.6 million nt (about 0.2% of the entire genome) from the NIH Intramural Sequencing Center at http://www.nisc.nih.gov/open_page.html?/projects/comp_seq.html. We developed a computer-based search profile in the C programming language that extracts all nonredundant sequences of 50 to 100 nt that are flanked by 10-nt to 30-nt perfect repeats. About 1,700 corresponding sequences could be extracted from the armadillo sequence information. Performing a RepeatMasker search (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker), we subsequently separated all unknown or tRNA-derived sequences for further processing. As most SINE elements are derived from tRNA genes, we also performed a local Blast search against all published tRNA-genes. In the end, we found about 20 copies of a tRNA-derived monomeric SINE element flanked by perfect direct repeats (DRs), an oligo(A) tail and an additional 5′ motif, GGGAA. These elements were locally blasted against all available *Dasypus novemcinctus* sequence data. With this strategy, we uncovered additional monomers, two types of dimers, and several types of chimeric sequences with tRNA and unknown sequence domains.

Characterization of the Novel SINE Elements

We aligned all extracted candidates of DAS-SINEs by using the Mac OS X/Darwin version of the MAFFT multisequence alignment program (Katoh et al. 2002). To improve the alignment, we performed local realignments with XCED, the graphical user interface of MAFFT. This alignment was used to distinguish several subtypes of DAS-SINEs. All DAS-SINEs feature a $tRNA^{Ala}$-like sequence, in most cases preceded by a short GGGAA-like motif. A secondary structure model of this element was drawn corresponding to secondary models compiled at the tRNAscan-SE Search Server (http://rna.wustl.edu/GtRDB/Hs/Hs-align.html) and mfold (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi) (default options; Zuker 2003) and compared with the human $tRNA^{Ala}$ structure that so far shows greatest similarity. Phylogenetic analysis was performed by the maximum-likelihood algorithm implemented in Tree-Puzzle version 5.2 (Schmidt et al. 2002) with the HKY model of substitution (Hasegawa, Kishino, and Yano 1985) and 1,000 quartet-puzzling steps (QPS). All DAS-SINE locations are assigned by the coordinates of GenBank entries (see Supplementary Material online).

Results

DAS-SINE Subfamilies

We discovered a novel high-copy-number family of SINE elements in the nine-banded armadillo that we term DAS-SINE (figs. 1 and 2). We extracted 687 copies of DAS-related members and could divide them into six subfamilies. Most DAS-SINE members are characterized by a 5′ extension, GGGAA, proximal to the tRNA region. The probably oldest members of DAS-SINEs are characterized by an oligo(A) tail classifying them as class T^- retroposons (Borodulina and Kramerov 2001). Compared with the typical oligo(A) tail, the remaining members terminate with a more complex A-rich consensus ($AATAAATCTTTA_{(n)}$) that assigns them to class T^+ SINEs. DAS-Ia has a monomeric structure with a tRNA core region. The tRNA region exhibits highest similarity to a human $tRNA^{Ala}$. At the class T^- branch (fig. 2*B*), a merger of two monomers separated by an oligo(A)-region forms a dimeric structure (DAS-IIa) akin to the 7SL-derived Alu dimers in primates. At the class T^+ branch, dimerization of two DAS monomers with more heterogeneous A-rich tails occurred, leading to the DAS-IIb subfamily. DAS-III chimeric elements are the result of a partial deletion of the second tRNA region of a DAS-IIb–like precursor that includes approximate 70 nt of the 5′ portion of the distal tRNA, as well as an insertion of approximately180 nt whose origin is currently unknown. This sequence shows no similarity to any GenBank entry. On the branch leading to DAS-III2b and DAS-III3b elements (fig. 2*B*), we found compensatory (U-A → U-G) and (U-A → C-G) changes that used to be the D-arm of the progenitor tRNA and represent a substitution profile unknown for any $tRNA^{Ala}$ D-arm. Concerning the master elements for DAS-SINEs, these changes (framed base pair) may indicate selection pressure toward a specific structure; for example, the one supported by the mfold secondary structure model of members of the DAS-Ia subfamily (fig. 1*B*). This structure occurs in the most stable mfold-predicted models of all DAS-SINE subfamily consensus sequences.

Finally, DAS-III3b elements exhibit an additional deletion of about 50 nt (see figure 2*B*). The 155 members of DAS-III3b can be further subdivided into a group of 135 members without additional insert, nine members with a specific 8-nt insertion, and 11 cases with a 19-nt insertion, indicating three different source genes (see Supplementary Material online).

Direct Repeats

A landmark of retroposed elements are their DRs. Perfect DRs are an indicator of recent retroposition events, and

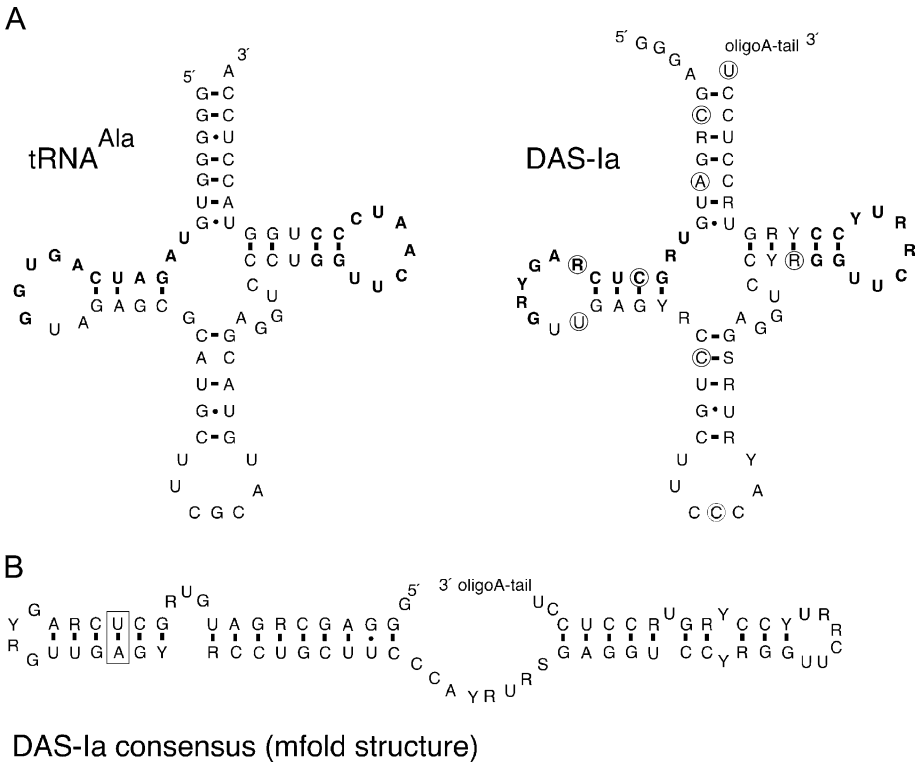

FIG. 1.—Secondary structure tRNA^{Ala} and structural similarities to the mature tRNA domain of the DAS-Ia consensus sequence. (*A*) Stucture of a tRNA (a human tRNA^{Ala} isoform) with close resemblance to the tRNA domain of DAS-SINEs. Internal promoter elements, box A and B, are shown in bold letters. Encircled letters represent major differences between the tRNA and the consensus sequence of the DAS-Ia SINE subtype. Dots denote G-U–base pairings. (*B*) The DAS-Ia consensus sequence is presented in the most frequently derived mfold structure of diverse individual sequences of DAS-Ia (as well as the other DAS-SINE) subfamily members. The boxed base pair represents the location of a prominent compensatory change in DAS-SINEs. Letters refer to the IUPAC code.

the existence of a retronuon subfamily with predominantly perfect DRs indicates recent activity of those elements. DAS-III3b retronuons exhibit the highest score of DR presence (90%) compared with DAS-II2b (50%) and DAS-Ia (44%) elements. This suggests that DAS-III elements are the youngest members of the DAS family (fig. 2*A*).

Self-Integration

A clear indicator of the evolutionary order of transpositions is the integration of SINE family members into already established older family members. We could detect DAS-III members in DAS-III and DAS-II types. On the other hand, we were unable to detect any DAS-I or DAS-II elements integrated into DAS-III retronuons (see Supplementary Material online). This provides additional evidence that DAS-III members are the most recent and/or still active elements. In that respect, they resemble Alu Y elements of primates (Batzer et al. 1996) (fig. 2*B*).

Phylogenetic Analysis

To compare the order of descent we established for the distinct DAS subfamilies, we performed a maximum-likelihood phylogenetic reconstruction of the tRNA-related parts (consensus sequences) of all DAS members. As an outgroup, we chose a human tRNA^{Ala}. Although the analyzed sequences are short, we could find strong correlation by high QPS values for the major groups of DAS members. DAS-III2b and DAS-III3b cluster by 100% QPS. All DAS-III members display one group with 100% QPS. The next relative is DAS-IIb, supported by 68% QPS. However, the basal resolution of the phylogenetic tree is too low to show a clear cluster of DAS-Ia and DAS-IIa. Furthermore, the origin of the second tRNA part of DAS-IIa and DAS-IIb cannot clearly be resolved by phylogenetic reconstruction (fig. 3).

Intraspecific Variability of DAS Subfamilies

We compared the variability of all 5′ terminal tRNA-derived sequence parts of DAS-SINE subfamilies expecting older subfamilies to be more heterogeneous than younger subfamilies. We could establish that DAS-Ia members are the most heterogeneous ($> 25\%$), whereas DAS-III3b members exhibit the lowest variability (5%) (figs. 4*A* and 5).

Target Site Preference

There are several options concerning genomic target-site selection of DAS elements: (1) random choice, (2) a consensus TTAAAA target preference, as suggested for the most abundant SINEs in rodents and primates (Jurka 1997), or (3) target sites complementary to the 3′ ends of the template RNAs, as shown for tailless retropseudogenes (Schmitz et al. 2004). We compared the nucleotide preference of target sites for the youngest and probably still

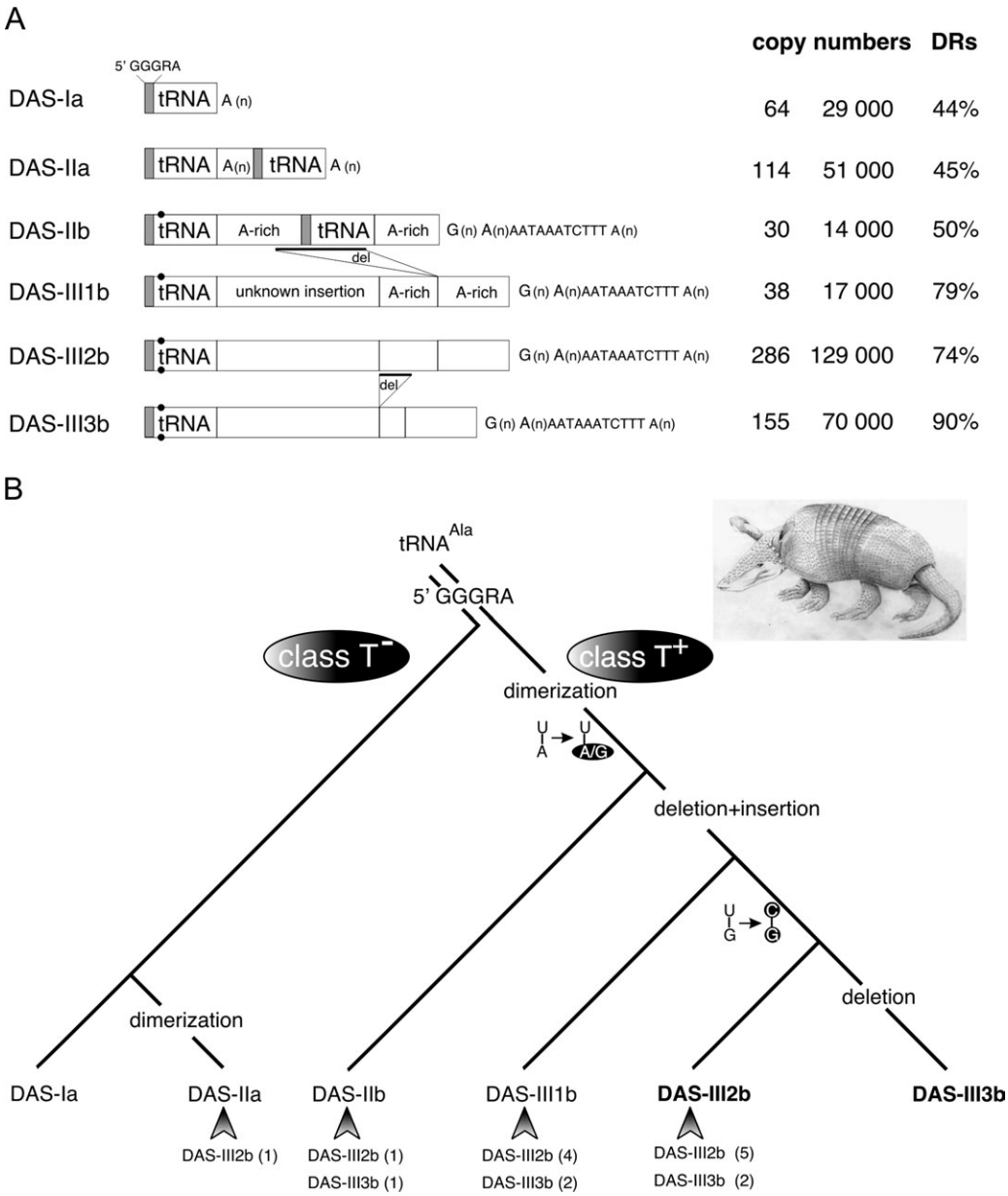

FIG. 2.—DAS-SINEs structures and evolution. (*A*) Schematic representation of DAS-SINE organization. Dots denote compensatory changes in structural domains. The numbers of analyzed DAS-SINEs subfamily members are shown with an estimation of total numbers in the *Dasypus* genome. Percentages denote the proportion of DAS-SINEs with recognizable direct repeats. (*B*) Hypothetical scenario for DAS-SINE subfamily evolution. The dimerizations, insertions, deletions, and compensatory changes are indicated on an evolutionary tree. The lower part of (*B*) represents retropositions of DAS elements into preexisting older DAS elements (quantity in parentheses).

active 354 DAS-III2b and DAS-III3b subfamily members and found a clear preference for a TTAAAA motif comparable to the SINEs described in Jurka (1997) (fig. 4*B*).

Discussion

The newly characterized tRNA related SINE family of the nine-banded armadillo is the first found in the edentates. According to Murphy et al. (2001), the edentates derive from the second branching point of the placentals, after splitting off from a common mammalian ancestor. Afrotherians were first to diverge and are characterized by a recently described SINE family (AfroSINEs [Nikaido et al. 2003]).

The novel *D. novemcinctus* SINE elements are currently not found in GenBank entries from any other species. Their taxonomic distribution within the edentates remains to be investigated experimentally. In total, we found 687 DAS-SINEs within the estimated 0.22% of genomic sequence information available (BAC clones). A total number of more than 300,000 copies per genome is assumed (fig. 2*A*). The DAS-SINE family is characterized by several clearly distinct subfamilies that probably originated from corresponding single or low-copy chromosomal source

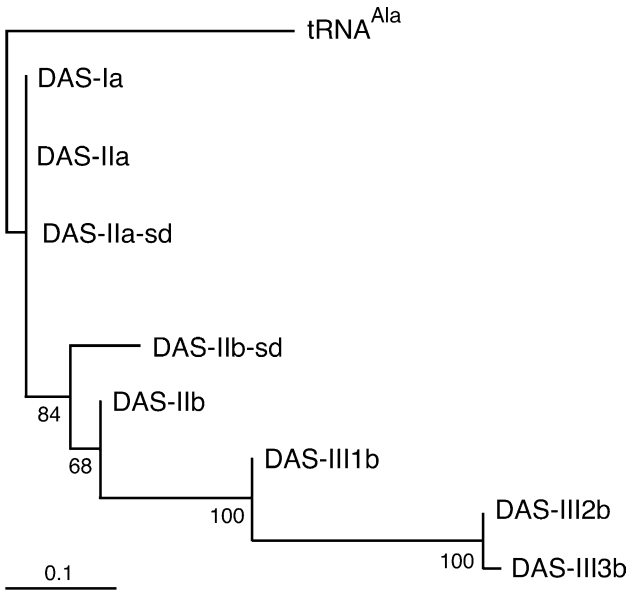

FIG. 3.—Phylogenetic reconstruction of the tRNA-related part of DAS-SINE consensus sequences. The human $tRNA^{Ala}$ sequence was used as an outgroup. Branch lengths represent nucleotide substitutions per site. Percentage QPS values are indicated at the corresponding nodes. Nodes supported by less than 60% were collapsed into polytomies.

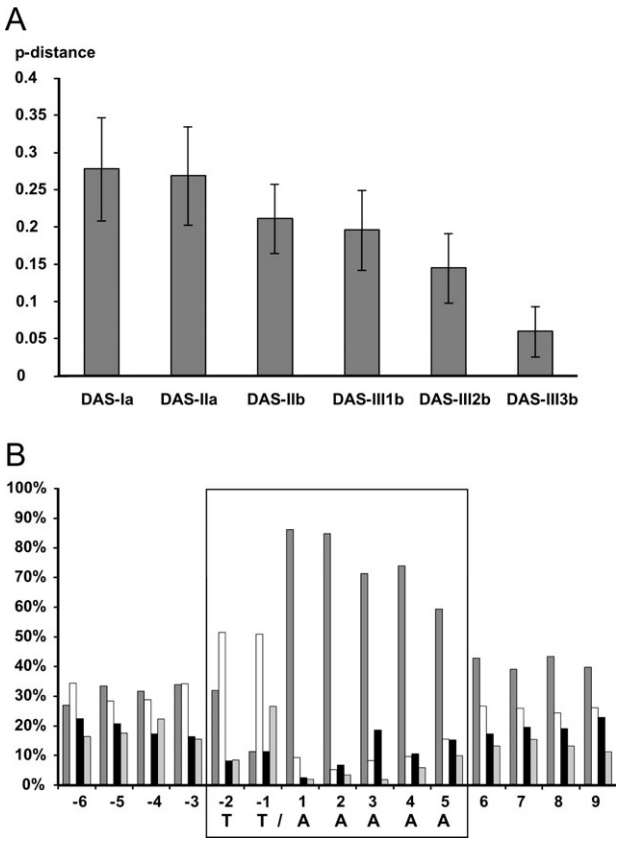

FIG. 4.—Sequence heterogeneity of DAS-SINE subfamilies and genomic target sites of the youngest DAS-III2 and DAS-III3 subfamilies. (*A*) *p*-Distances indicate a correlation between age and heterogeneity within subfamilies. Standard errors are indicated. (*B*) Genomic target-site preferences for a TT/AAAAA motif. The backslash represents the genomic cutting site. Dark-gray bars represent adenine, white bars represent thymine, black bars represent guanine, and light-gray bars represent cytosine.

genes that ''seeded'' additional SINE source genes into different genomic locations. A high degree of similarity to a human $tRNA^{Ala}$ gene and the distinct tRNA like structural features (figs. 1 and 5) suggest $tRNA^{Ala}$ as the initial source for DAS-SINEs. An additional common characteristic of DAS elements is a GGGAA-like motif at their 5′ ends that presumably corresponds to the 5′ tRNA flanking region of the original chromosomal locus. Most tRNA-derived SINEs start with similar G-rich and 3 to 7 nt long tRNA-unrelated 5′ regions, possibly because of the use of a partially unprocessed tRNA as retroposition template. An example is the rodent-specific BC1 RNA that probably arose by retroposition of a partialy processed $tRNA^{Ala}$ with a 5′ GCGGCT leader sequence (Rozhdestvensky et al. 2001). Correspondingly, the BC1 leader sequence starts with a similar 5′ GGGGTT motif. Mature tRNAs, without any leader sequence, may serve as primers in reverse transcription of certain retroviruses and long terminal repeat (LTR) retrotransposons (Mak and Kleiman 1997). They may also serve as templates for the generation of certain SINE elements (Ohshima et al. 1996). DAS-SINEs are clearly distinct from this mechanism because of their unprocessed state.

A possible sequence of events forming the distinct DAS subfamilies can be summarized as follows: A first wave of retroposition generated a monomeric subfamily of DAS-Ia SINEs with a short leader sequence, characteristic homogeneous oligo(A) tail, and flanking DRs. The dimeric DAS-IIa retronuons probably arose by a DAS-Ia integration into the oligo(A) tail of a preexisting DAS-Ia master gene and subsequent transcription and retroposition of the dimeric DAS-IIa elements, featuring two tRNA domains separated by an oligo(A) region.

Perhaps in an independent second wave, as indicated by parallel lines in figure 2*B*, a more complex dimeric group of DAS-SINEs arose featuring additional approximately 120 nt long, heterogeneous A-rich motifs distal to the tRNA parts. Presumably, the A-rich tail of a monomeric DAS master gene grew in size substantially, presumably by extension during the retroposition process. This phenomenon has been described for Alu elements by Hagan, Sheffield, and Rudin (2003) and Dewannieux, Esnault, and Heidmann (2003) and is thought to be the result of reverse transcriptase slippage. During dimerization, the corresponding monomeric transcript retroposed into the enlarged 3′ flank of a corresponding master gene to yield the dimeric structure of DAS-IIb. Analysis of DAS-SINE retropositions into preexisting DAS-SINE loci suggests that in 14 out of 16 cases, the integration took place in the A-rich spacer region between the two tRNA domains. This indicates a high acceptor affinity of those regions, probably a situation comparable to the formation of the first dimeric structures. However, we could not detect the typical TTAAAA target motif.

DAS-III chimeric elements simply emerged because of a deletion and insertion within a DAS-IIb–like master locus rather than by template switching. Hence, scenario two (see *Introduction*) in combination with insertions and deletions are the most probable events in forming dimeric and chimeric DAS subfamilies. In contrast to DAS-Ia and

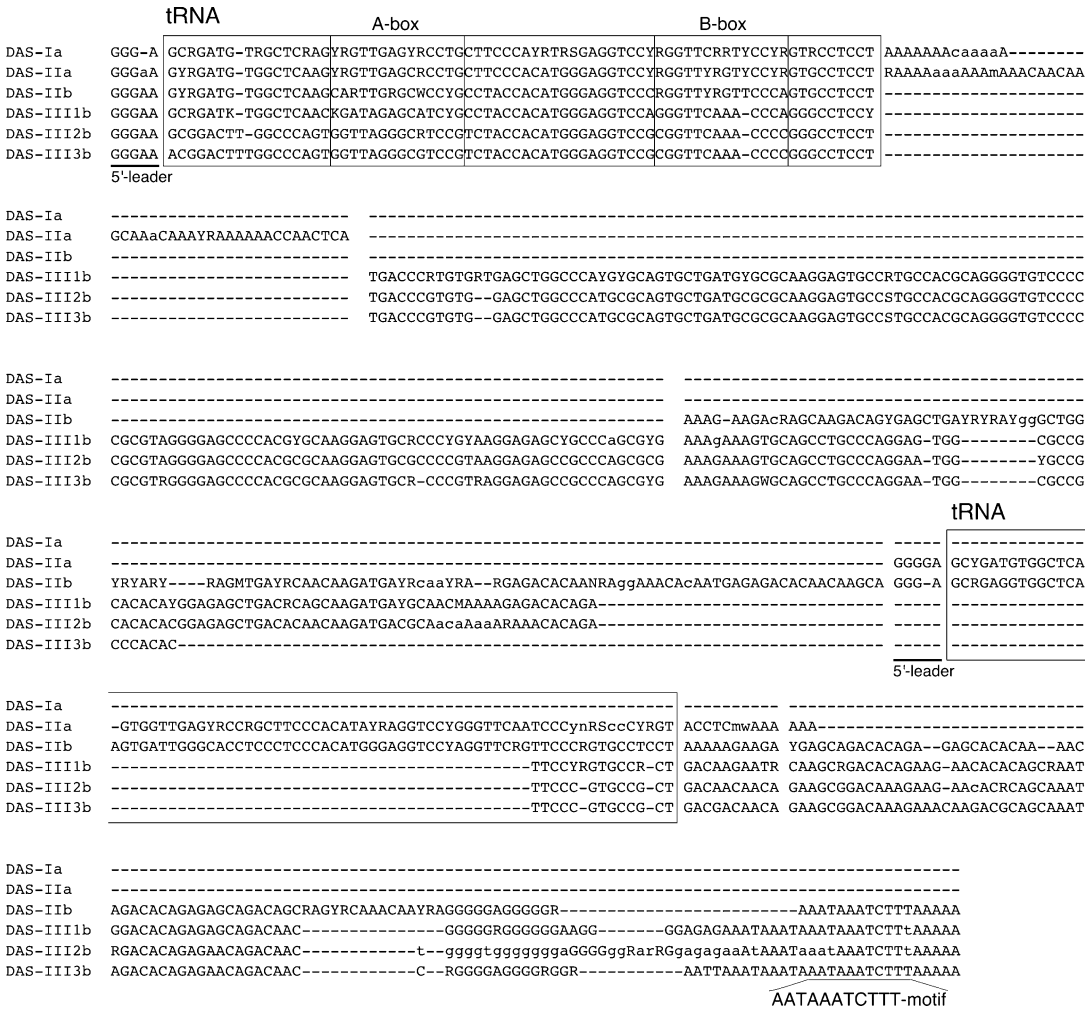

FIG. 5.—Consensus sequences and organization of DAS-SINEs. The 5′ leader sequences are underlined, and the tRNA like domains are boxed. The class T^+ diagnostic motif is shown at the 3′ ends of the DAS-IIb and DAS-III subfamilies sequences.

DAS-IIa, all DAS-IIb and DAS-III subfamilies terminate with a 3′ ATAAATCTTT-like motif, the trait of class T^+ elements (Borodulina and Kramerov 2001). This observation supports the hypothesis of an independent origin of DAS-Ia /DAS-IIa from DAS-IIb/DAS-III subfamilies, even though both are derived from similar RNA template.

The DAS-III3b subfamily is the youngest, with 90% of the members exhibiting distinct DRs, as well as the lowest degree of internal heterogeneity (see figure 5). Finally, DAS-III3 elements are found integrated in most other DAS members but not vice versa, indicating a clear pattern of appearance over time (see figure 2*B*). Most of the splitting points as drawn by key events of deletions or insertions are congruent with the sequence phylogeny of the tRNA-related consensus part of all DAS-SINE subfamilies. Although the phylogenetic reconstruction was restricted to the tRNA related parts, strong QPS support values could be derived for the DAS-IIb and DAS-III subfamilies. For determination of a close relationship between DAS-Ia and DAS-IIa, as well as the affiliation of the second tRNA related region of DAS-IIa and DAS-IIb, sequence phylogeny could not be applied (fig. 3). This underscores once more the superiority of insertion and deletion analysis for phylogenetic inquiry.

In the lineage leading to the class T^+ DAS-SINEs compensatory changes occurred (U-A → U-G → C-G [see figure 2]) at the RNA level. These changes correspond to the D-arm of the $tRNA^{Ala}$ and are unknown for any published tRNA. Consequently, the DAS-specific compensatory change seems to be under structural selective pressure to conserve a tRNA D-arm–like structure in the founder RNAs (see figure 1*B*). This indicates that the respective master genes may encode a functional RNA.

For 354 DAS-III SINEs with recognizable DRs, we reconstructed the nucleotide composition of the genomic target sites and derived a TTAAAAA consensus sequence (fig. 5*B*; boxed area). SINE retroposition depends on the transpositional machinery of autonomous elements, such as LINEs. For Alu-SINEs, L1-mediated reverse transcription and integration is the most apparent mechanism of retroposition (Jurka 1997). Both retronuons share a characteristic 3′ end, oligo(A), that is responsible for their target-site preference. Concerning the target-site preference, DAS-SINEs show the same integration profile as Alu-SINEs

and as a consequence, share similar 3′ ends. In contrast, L2 or L3 and their nonautonomous associates deviate substantially (Kapitonov, Pavlicek, and Jurka 2004). Presumably, L2 and L3 elements were active 200 to 300 MYA, long before the mammalian radiation (Kapitonov, Pavlicek, and Jurka 2004). To selectively detect LINE elements in armadillo whose activity coincides with that of DAS-SINEs, we extracted all LINE-related sequences flanked by perfect DRs (data not shown). Together with the aforementioned target-site preference, detection of L1 elements only supports our notion that L1 elements mediated retroposition of DAS-SINE elements.

In conclusion, we could follow the evolution of a newly discovered SINE family in edentates, the first described in this order to date. The origin can clearly be traced back to a presumably incompletely processed $tRNA^{Ala}$ yielding a monomeric SINE master gene, followed by the emergence of dimeric and chimeric forms. Compensatory changes point to a structural constraint on the tRNA D-arm-corresponding region of the DAS-SINEs source gene. Other features correspond to canonical characteristics of SINEs (e.g., flanking DRs, oligo(A) terminal regions, and the preference for AT-rich target sites). This preference is shared by resident armadillo L1 elements that probably were active at the same period and provided the necessary enzymatic machinery for DAS-SINE retroposition. Our study completes evidence that SINE elements are present in all mammalian orders. It will facilitate further investigations of SINE elements and their evolutionary impact in the order Xenathra.

Acknowledgments

We thank Timofey Rozhdestvensky, Christian Roos, and Silke Singer for helpful comments on the manuscript and Tomas Schmitz for the drawing of an armadillo. Many thanks go to Kira Gee for revising the English text. The work was supported by the German Human Genome Project through the BMBF (01KW9966) to J.B.

Literature Cited

Batzer, M. A., P. L. Deininger, U. Hellmann-Blumberg, J. Jurka, D. Labuda, C. M. Rubin, C. W. Schmid, E. Zietkiewicz, and E. Zuckerkandl. 1996. Standardized nomenclature for Alu repeats. J. Mol. Evol. **42**:3–6.

Bentolila, S., J. M. Bach, J. L. Kessler, I. Bordelais, C. Cruaud, J. Weissenbach, and J. J. Panthier. 1999. Analysis of major repetitive DNA sequences in the dog (*Canis familiaris*) genome. Mamm. Genome **10**:699–705.

Borodulina, O. R., and D. A. Kramerov. 2001. Short interspersed elements (SINEs) from insectivores: two classes of mammalian SINEs distinguished by A-rich tail structure. Mamm. Genome **12**:779–786.

Brosius, J. 1999. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene **238**:115–134.

Brosius, J., and S. J. Gould. 1992. On ''genomenclature'': a comprehensive (and respectful) taxonomy for pseudogenes and other ''junk DNA''. Proc. Natl. Acad. Sci. USA **89**:10706–10710.

Buzdin, A. A. 2004. Retroelements and formation of chimeric retrogenes. Cell. Mol. Life Sci. **61**:1–14.

Cheng, J. F., R. Printz, T. Callaghan, D. Shuey, and R. C. Hardison. 1984. The rabbit C family of short, interspersed repeats: nucleotide sequence determination and transcriptional analysis. J. Mol. Biol. **176**:1–20.

Cost, G. J., Q. Feng, A. Jacquier, and J. D. Boeke. 2002. Human L1 element target-primed reverse transcription in vitro. EMBO J. **21**:5899–5910.

Daniels, G. R., and P. L. Deininger. 1991. Characterization of a third major SINE family of repetitive sequences in the galago genome. Nucleic Acids Res. **19**:1649–1656.

Dewannieux, M., C. Esnault, and T. Heidmann. 2003. LINE-mediated retrotransposition of marked Alu sequences. Nat. Genet. **35**:41–48.

Feschotte, C., N. Fourrier, I. Desmons, and C. Mouches. 2001. Birth of a retroposon: the twin sine family from the vector mosquito *Culex pipiens* may have originated from a dimeric tRNA precursor. Mol. Biol. Evol. **18**:74–84.

Gilbert, N., and D. Labuda. 2000. Evolutionary inventions and continuity of CORE-SINEs in mammals. J. Mol. Biol. **298**:365–377.

Hagan, R., F. Sheffield, and C. M. Rudin. 2003. Human Alu element retroposition induced by genotoxic stress. Nat. Genet. **35**:219–220.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22**:160–174.

Jurka, J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc. Natl. Acad. Sci. USA **94**:1872–1877.

Kajikawa M., and N. Okada. 2002. LINEs mobilize SINEs in the eel through a shared 3′ sequence. Cell **111**:433–444.

Kapitonov, V. V., and J. Jurka. 2003. A novel class of SINE elements derived from 5S rRNA. Mol. Biol. Evol. **20**:694–702.

Kapitonov, V. V., A. Pavlicek, and J. Jurka. 2004. Anthology of human repetitive DNA. Pp. 251–306 *in* R. A. Meyers, ed. Encyclopedia of molecular cell biology and molecular medicine, Vol. 1. Wiley VCH, Weinheim, Germany,.

Kass, D. H., J. Kim, and P. L. Deininger. 1996. Sporadic amplification of ID elements in rodents. J. Mol. Evol. **42**:7–14.

Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. **30**:3059–3066.

Kramerov, D. A., A. A. Grigoryan, A. P. Ryskov, and G. P. Georgiev. 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.

Lander, E. S., L. M. Linton, B. Birren, et al. (256 co-authors) 2001. Initial sequencing and analysis of the human genome. Nature **409**:860–921.

Lin, Z., O. Nomura, T. Hayashi, Y. Wada, and H. Yasue. 2001. Characterization of a SINE species from vicuna and its distribution in animal species including the family Camelidae. Mamm. Genome **12**:305–308.

Mak, J., and L. Kleiman. 1997. Primer tRNAs for reverse transcription. J. Virol. **71**:8087–8095.

Murphy, W. J., E. Eizirik, W. E. Johnson, Y. P. Zhang, O. A. Ryder, and S. J. O'Brien. 2001. Molecular phylogenetics and the origins of placental mammals. Nature **409**:614–618.

Nikaido, M., H. Nishihara, Y. Hukumoto, and N. Okada. 2003. Ancient SINEs from African endemic mammals. Mol. Biol. Evol. **20**:522–527.

Nikaido, M., A. P. Rooney, and N. Okada. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. USA **96**:10261–10266.

Nishihara, H., Y. Terai, and N. Okada. 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.

Ohshima, K., M. Hamada, Y. Terai, and N. Okada. 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.

Okada, N., and K. Ohshima. 1995. Evolution of tRNA-derived SINEs. Pp. 61–79 *in* R. J. Maraia, ed. The impact of short interspersed elements (SINEs) on the host genome. RG Landes Company, Austin, Tex.

Piskurek, O., M. Nikaido, Boeadi, M. Baba, and N. Okada. 2003. Unique mammalian tRNA-derived repetitive elements in Dermopterans: The t-SINE family and its retrotransposition through multiple sources. Mol. Biol. Evol. **20**:1659–1668.

Quentin, Y. 1992. Fusion of a free left Alu monomer and a free right Alu monomer at the origin of the Alu family in the primate genomes. Nucleic Acids Res. **20**:487–493.

Rogers, J. H. 1985. The origin and evolution of retroposons. Int. Rev. Cytol. **93**:187–279.

Roos, C., J. Schmitz, and H. Zischler. 2004. Primate jumping genes elucidate strepsirrhine phylogeny. Proc. Natl. Acad. Sci. USA **101**:10650–10654.

Rozhdestvensky, T. S., A. M. Kopylov, J. Brosius, and A. Huttenhofer. 2001. Neuronal BC1 RNA structure: evolutionary conversion of a tRNA(Ala) domain into an extended stem-loop structure. RNA **7**:722–730.

Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics **18**:502–504.

Schmitz, J., G. Churakov, H. Zischler, and J. Brosius. 2004. A novel class of mammalian-specific tailless retropseudogenes. Genome Res. **14**:1911–1915.

Schmitz, J., and H. Zischler. 2003. A novel family of tRNA-derived SINEs in the colugo and two new retrotransposable markers separating dermopterans from primates. Mol. Phylogenet. Evol. **28**:341–349.

Shimamura, M., H. Yasue, K. Ohshima, H. Abe, H. Kato, T. Kishiro, M. Goto, I. Munechika, and N. Okada. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature **388**:666–670.

Szemraj, J., G. Pluciennicczak, J. Jaworski, and A. Pluciennicczak. 1995. Bovine Alu-like sequences mediate transposition of a new site-specific retroelement. Gene **152**:261–264.

Ullu, E., and C. Tschudi. 1984. Alu sequences are processed 7SL RNA genes. Nature **312**:171–172.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31**:3406–3415.

Pierre Capy, Associate Editor

Accepted December 20, 2004