A young *Alu* subfamily amplified independently in human and African great apes lineages

Ewa Ziętkiewicz*, Chantal Richer, Wojciech Makalowski+, Jerzy Jurka[1] and Damian Labuda
Centre de Recherche, Hôpital Ste-Justine, Département de Pédiatrie, Université de Montréal, Montréal, Québec, H3T1C5, Canada and [1]Linus Pauling Institute of Science and Medicine, Palo Alto, CA 94306, USA

Received September 20, 1994; Revised and Accepted October 26, 1994 GenBank accession nos: U14684–U14719 (incl.)

SUMMARY

A variety of *Alu* subfamilies amplified in primate genomes at different evolutionary time periods. *Alu* Sb2 belongs to a group of young subfamilies with a characteristic two-nucleotide deletion at positions 65/66. It consists of repeats having a 7-nucleotide duplication of a sequence segment involving positions 246 through 252. The presence of Sb2 inserts was examined in five genomic loci in 120 human DNA samples as well as in DNAs of higher primates. The lack of the insertional polymorphism seen at four human loci and the absence of orthologous inserts in apes indicated that the examined repeats retroposed early in the human lineage, but following the divergence of great apes. On the other hand, similar analysis of the fifth locus (butyrylcholinesterase gene) suggested contemporary retropositional activity of this subfamily. By a semi-quantitative PCR, using a primer pair specific for Sb2 repeats, we estimated their copy number at about 1500 per human haploid genome; the corresponding numbers in chimpanzee and gorilla were two orders of magnitude lower, while in orangutan and gibbon the presence of Sb2 *Alu* was hardly detectable. Sequence analysis of PCR-amplified Sb2 repeats from human and African great apes is consistent with the model in which the founding of Sb2 subfamily variants occurred independently in chimpanzee, gorilla and human lineages.

INTRODUCTION

Ubiquitous short interspersed repeats (SINEs) are a hallmark of mammalian genomes. They spread in great numbers by retroposition, i.e. by genomic reintegration of their RNA transcripts (for a recent review see [1]). *Alu* elements which amplified in the primate lineage belong to 7SL RNA-derived retroposons. Represented by more than half a million copies they contribute about 5% to the bulk of human genomic DNA. A number of *Alu* subfamilies have spread at different periods of primate evolution through retropositionally active sequences that can be now reconstructed as consensus sequences for these subfamilies. The highest degree of divergence among the repeats, and the closest resemblance of the corresponding consensus to the 7SL RNA sequence are seen in the oldest subfamilies. The average age of the oldest subfamily of *Alu* J repeats was estimated at 55 Myr using the molecular clock [2], suggesting that the peak of their amplification occurred during the Paleocene period. The greatest genomic expansion of *Alu* repeats took place prior to simian radiation; it included the subfamilies Sx, Sp and Sq, formerly grouped as the subfamily Sa of an estimated age of 31 Myr (see [3] for *Alu* classification, and references therein). A group of younger human repeats, represented by human subfamilies Sc and Sb (age of 24 and 18 Myr, respectively [2]), appeared during a slow-down period of *Alu* dispersal. These young *Alu* subfamilies are marked by a characteristic two-nucleotide deletion at positions 65/66 and can be found in all simian lineages [4]. Besides little sequence divergence among member repeats, young *Alu* subfamilies display a considerable extent of species-specificity as well as intra-species insertional polymorphism. Most of the recent *Alu* inserts found in one primate lineage are absent from orthologous positions in closely related species, because their retroposition occurred following speciation [5–12]. Within species, many of the insertions are polymorphic and some of them are found in only few individuals, thus reflecting dispersion events that occurred in the near past [7–10,12–18].

The potential to identify and characterize genomic loci that are at the origin of *Alu* transcripts retroposing in contemporary genomes, makes studying small, recently amplifying *Alu* subfamilies especially attractive. From such investigations we can learn about the dispersal dynamics of *Alu* subfamilies: whether more than one subfamily is active during the same period and in the same genome, and/or whether more than one locus transcribes *Alu* RNA undergoing retroposition. Structural constraints on actively retroposing *Alu* repeats can be eventually defined and the active *Alu* species may be also cloned and used in further experiments. For these reasons we undertook

*To whom correspondence should be addressed

+Present address: National Center for Biotechnology Information, National Library of Medicine, NIH, 8600 Rockville Pike, Bethesda, MD 20894, USA characterization of the recently described [18,19] *Alu* subfamily Sb2. This subfamily has a characteristic 7-nucleotide duplication involving positions 246 through 252. Its GenBank density suggested no more than 1000−2000 copies in the human genome. The presence of an Sb2 insert in the vicinity of the Huntington disease (HD) gene, seen only in two families affected with this disease, and its absence from more than a thousand unrelated chromosomes implied its recent retropositional activity [18]. Low sequence divergence among the few known Sb2 repeats and their similarity to other young *Alu* subfamilies was also consistent with their recent dispersal [18]. To trace the origin of the Sb2 subfamily and to measure its amplification, we investigated intraspecies and inter-species polymorphisms resulting from Sb2 insertions, and examined the frequency of these repeats in human and related primate genomes. Our data indicate that Sb2 subfamily started to disperse early in the human lineage, while its distinct variants amplified, albeit to a much lower extent, in chimpanzee and gorilla.

MATERIALS AND METHODS

PCR amplification of Sb2-containing loci was performed using primer pairs (0.5 μM each) flanking the targeted Sb2 insertions: 5′−GCAAAGACTTCCCAGAG−3′ and 5′− GGGTCCTCTCAAACTTCTG−3′ (for lecithin−cholesterol acyltransferase, LCAT, locus); 5′− AGCAACTACCCATCCA−3′ and 5′−CACGACAGCCTCAAATG−3′ (low density lipoprotein receptor, LDLR); 5′−CAGTAGGATGTCACCATATC−3′ and 5′−AGGGACAGGGAAAGATG− 3′ (biliary glycoprotein, BGP); 5′−GCCCAGCCAATTTAATTAT−3′ and 5′− GCTGCATTGGAGCAAATA−3′ (oestrogen receptor, OESR); 5′− CCACGAGGACCGAAGTC−3′ and 5′−GCCTCACGGTAGTTTTCAG−3′ (butyrylcholinesterase, CHEB); at annealing temperatures of 56°C, 52°C, 56°C, 52°C and 57°C, respectively. The PCR reaction conditions were as described [20], except that 4% formamide was included in the buffer and that 30 PCR cycles were performed, each consisting of 30 s at 94°C, 30 s at the annealing temperature (as given above) and 30 s at 72°C. PCR products were fractionated by electrophoresis in 1% agarose gel, transferred onto a nylon membrane (Hybond/Amersham) and hybridized with [^{32}P]-labelled Sb2-specific oligonucleotide, INS2, overlapping the 7-bp duplication (5′-GGCCGGACTGCGGACTGCA-3′). The hybridization was carried out in 1 M NaCl, 1% SDS and 5× Denhardt at 64°C for 2 h, followed by two washes at room temperature for 10−15 min. and one wash at 64°C for 20 min, in 2×SSC/0.1% SDS. Targeted amplification of Sb2 repeats was performed at annealing temperature of 62°C, using primers: R16A/6 (5′−GGCGCGGTGGCTCACG−3′) corresponding to the 5′-end of the *Alu* consensus and the Sb2-specific oligonucleotide INS2 (see above and Fig. 1). Radioactive α−[^{32}P]dCTP (Amersham 800 Ci/mmol) was included in the reaction at the final concentration of 0.125 μM. Following separation by electrophoresis in 6% non-denaturing polyacrylamide gel [21], the amplification products were excised from the dried gels and quantified by Cerenkov counting. The genomic copy number was estimated from a calibration curve based on [^{32}P]-counts obtained from linearized, Sb2-containing pBS plasmid used as a copy number standard. The amplification products obtained using the same primer pair or, alternatively, R16A/6 and L23A/266 instead of INS2 (5′−TTTTTTGAGACGGAGTCTCGCTC−3′, see Fig. 1) were fractionated by electrophoresis as above. They were eluted from the gel slices, cloned in pBS plasmid and sequenced [20,21].

DNA samples were as described in [22,23], except for the pygmy chimpanzee (*Pan paniscus*) DNA obtained from Jurgen Brosius.

RESULTS

To test whether Sb2 insertions were polymorphic or fixed in the human population, we screened 120 human DNA samples of Caucasian, Oriental and Black origin at five genomic loci where Sb2 repeats were previously detected [18,19]. At four such locations, lecithin-cholesterol acyltransferase (LCAT), biliary glycoprotein (BGP), low density lipoprotein receptor (LDLR) and oestrogen receptor (OESR) genes, Sb2 inserts were present in all human samples analyzed using PCR technique (shown in Fig. 2 and summarized in Table 1). The presence of Sb2 inserts was inferred from the size of the amplification products and was confirmed by hybridization using the INS2 oligonucleotide as an Sb2-specific probe [note that in the case of the LDLR locus (Fig. 2), due to a mismatch with the INS2 probe, the hybridization

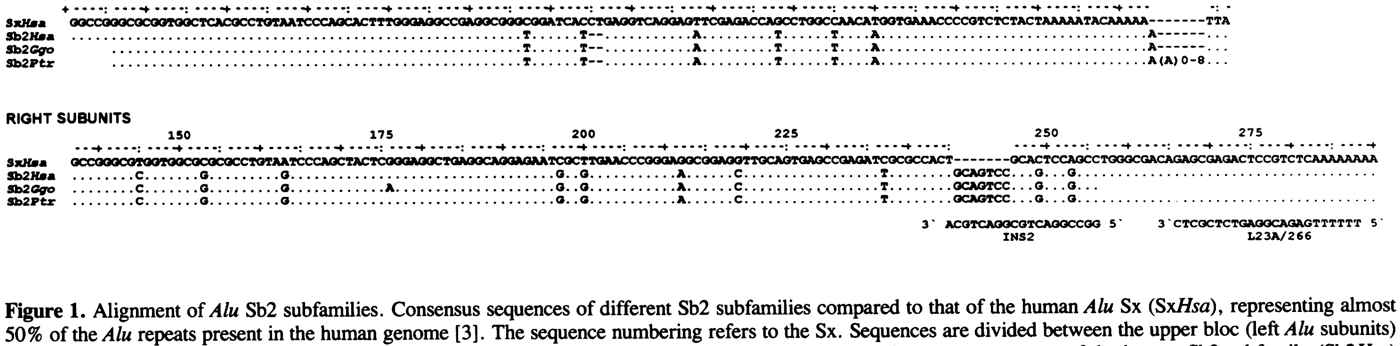
LEFT SUBUNITS
R16A/6
3`GGCGCGGTGGCTCACG 3`
0
25
50
75
100
125
SxHsa GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAA-------TTA
Sb2Hsa
Sb2Ggo
Sb2Ptr
A(A)0-8
RIGHT SUBUNITS
150
175
200
225
250
275
SxHsa GCCGGGCGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCGAGATCGCGCCACT-------GCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAAAAA
Sb2Hsa
Sb2Ggo
Sb2Ptr
GCAGTCC
3` ACGTCAGGCGTCAGGCCGG 5`
INS2
3`CTCGCTCTGAGGCAGAGTTTTTT 5`
L23A/266

Figure 1. Alignment of *Alu* Sb2 subfamilies. Consensus sequences of different Sb2 subfamilies compared to that of the human *Alu* Sx (Sx*Hsa*), representing almost 50% of the *Alu* repeats present in the human genome [3]. The sequence numbering refers to the Sx. Sequences are divided between the upper bloc (left *Alu* subunits) and the lower bloc (right subunits). Dots denote identity with the top sequence, dashes represent gaps. Consensus sequence of the human Sb2 subfamily (Sb2*Hsa*) was derived from 24 sequences including nine from the GenBank release 84.0 (i.e. from the five loci analyzed in this paper, the HD locus, HSSTREP, HDAB and EST06329). Consensus of gorilla (Sb2*Ggo*) and that of chimpanzee (Sb2*Ptr*) were derived from 11 and 10 sequences, respectively, and are truncated at the priming sites used. The sequences of oligonucleotides used as PCR primers (R16A/6, INS2) and a hybridization probe (INS2) are indicated. '(A)0−8' in the chimpanzee consensus indicates the linker length polymorphism within this subfamily.

was only detected at a relaxed stringency (not shown)]. These four Sb2 inserts appear thus to be fixed in the human population, indicating relatively ancient retropositional events. However, their absence from the orthologous sites in apes (Fig. 2) indicates retropositions following the divergence of human and great apes (note that the results for the LDLR locus were inconclusive here

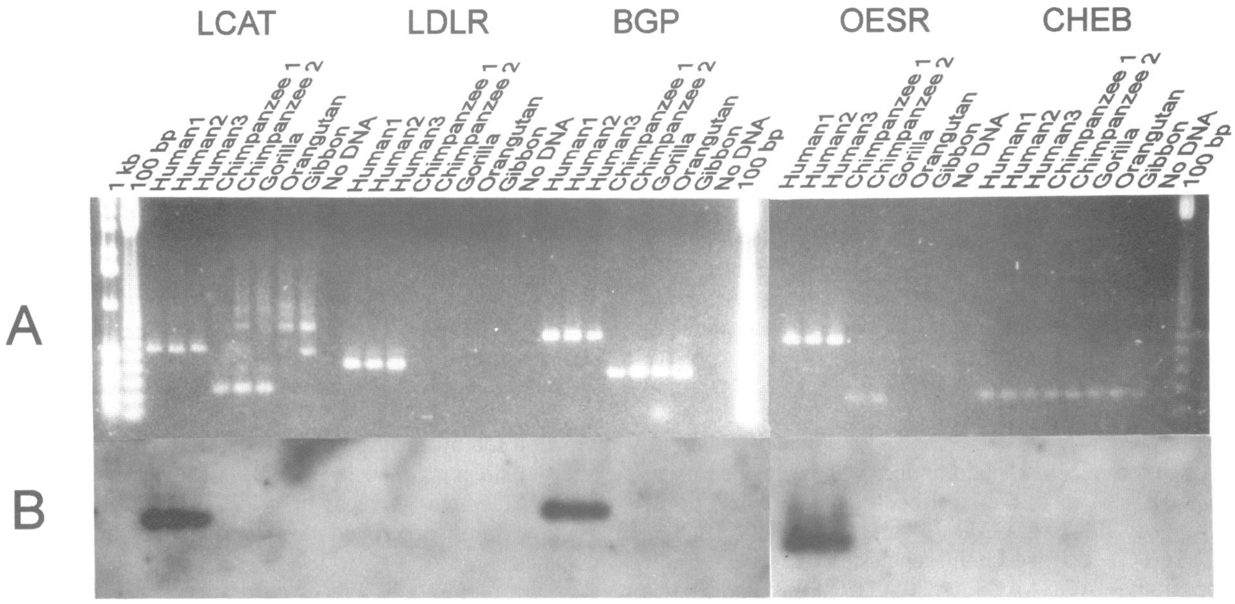

Figure 2. PCR analysis of five genomic loci where Sb2 inserts have been shown to occur [18,19]. Origin of DNA samples is indicated at the top. The expected lengths of Sb2-containing fragments amplified from LCAT, LDLR, BGP, OESR and CHEB were 564, 444, 863, 470 and 509 bp, respectively. Ethidium bromide-stained agarose gel and the hybridization of the amplified genomic segments with [^{32}P]-labelled Sb2-specific oligonucleotide probe INS2 are shown in (A) and in (B), respectively. A single mismatch between this probe and Sb2 target in the human LDLR locus explains lack of hybridization in this case.

Table 1. Summary of PCR and hybridization results at five loci reported to contain Sb2 insert

Locus	PCR amplification of Sb2-containing loci									INS2-oligo hybridization
	Base pairs*	Human			Apes and monkeys					
		Orient. (68)	Blacks (32)	Caucas. (20)	Chimp. (2)	Gorilla (1)	Orang. (1)	Gibbon (1)	Baboon (1)	Human (3)
LCAT	564	1	1	1	0	0	0	0	0	1
	252	0	0	0	1	1	0	0	0	
LDLR	444	1	1	1	0	0	0	0	0	1**
	119	0	0	0	1	1	0	0	0	
BGP	863	1	1	1	0	0	0	0	0	1
	440	0	0	0	1	1	1	0	0	
OESR	470	1	1	1	0	0	0	0	0	1
	339	0	0	0	1	0	0	0	0	
CHEB	509	0	0	0	0	0	0	0	0	0
	167	1	1	1	1	1	1	1	0	

The number of individual DNA samples tested is indicated between parentheses.
*expected length of the amplified fragment in the presence (upper/shaded) and absence (lower) of the Sb2 insert.
**seen only at relaxed conditions.

since the human-specific primers we used failed to direct the amplification with non-human templates).

In contrast to the Sb2 inserts in four genomic loci described above, that in the butyrylcholinesterase (CHEB) gene causes disruption of the open reading frame [15]. We did not find this particular Sb2 in any of the human or ape samples analyzed, indicating that the reported insert [15] represented a recent retropositional event, similar to that in the HD locus [18]. These two insertions are thus consistent with the Sb2 subfamily being still able to generate new retropositions.

Semi-quantitative PCR, directed at Sb2 repeats, was used to determine whether the dispersal of Sb2 *Alu* subfamily was restricted to the human lineage or whether it also occurred in the related primates. We selectively targeted *Alu* Sb2 repeats by using an R16A/6 oligonucleotide as a 5'-*Alu* primer (see [21] for the nomenclature of *Alu*-specific PCR primers) and the INS2 oligonucleotide as an Sb2-specific 3'-primer. It was necessary to vary the initial amount of a genomic template and the number of PCR cycles to allow for simultaneous observation of the amplification of Sb2-like *Alu* repeats from human and apes (see Fig. 3). While a single prominent band of the expected size was amplified in human and gorilla DNAs, the PCR products in chimpanzee consisted of a smear of bands differing by few base pairs. In orangutan and gibbon PCR products of the expected size were barely detectable, indicating the presence of no more than few, if any, genomic copies of *Alu* repeats with the 7-nucleotide duplication. Using an Sb2-clone (linearized plasmid) as a standard template, we estimated the number of Sb2 repeats in the human genome at approximately 1500–2000 copies per haploid equivalent, consistent with the expectation from the GenBank data. The corresponding numbers in chimpanzee and gorilla were two orders of magnitude lower (15–20 and 15–25 copies, respectively). Although the copy number determination using this approach has to be taken with caution (estimated accuracy not higher than 50%), it reflects the relative abundance of Sb2 repeats in the human and great apes genomes.

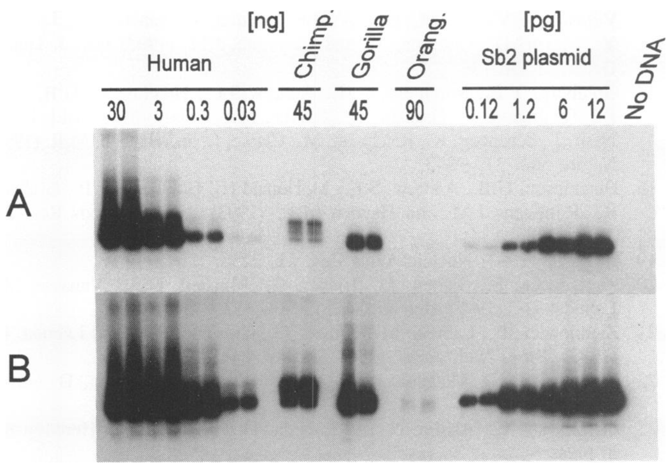

Figure 3. Quantification of Sb2 repeats by PCR. Amplification products after 20 (A) and 24 (B) PCR cycles were resolved by polyacrylamide gel electrophoresis. The origin and the amount of DNA template is indicated at the top of the autoradiograms.

To confirm the identity of the quantified PCR-products, the corresponding bands were extracted from the gel, cloned and sequenced. We determined 15 human, 11 chimpanzee and 10 gorilla sequences. The derived consensus sequences are reported in Figure 1 (the human consensus is based on the alignment of 24 sequences, including nine repeats from the GenBank release 84.0). The gorilla (*Gorilla gorilla*) Sb2 consensus sequence differs from the human one by a single substitution (G→A) at position 176. Although the common chimpanzee (*Pan troglodytes*) sequence is identical with the human Sb2 in both subunits, the A-rich linker (with usually five A-residues between positions 128 and 133 in the general *Alu* consensus) is expanded up to fourteen A residues, explaining the presence of non-homogenous PCR products seen in Figure 3. Similar length variation is present in a pygmy chimpanzee (*Pan paniscus*) as well (not shown). However, the electrophoretic patterns of Sb2 repeats PCR-amplified from the common and pygmy chimpanzee DNAs are not identical, indicating that Sb2 subfamily was retropositionally active in these two species following their divergence.

Based on the sequence dissimilarity from the consensus (Table 2) at non-CpG positions and assuming a mutation rate of 0.15% per million years [24], we estimate the average age of Sb2 subfamilies in the human, gorilla and chimpanzee lineages at 5.0, 4.3 and 2.5 Myr), respectively. These are consistent with the previous estimation of the human Sb2 average age at 4.1 Myr, based on the data from six Sb2 repeats [18]. The corresponding estimations (Table 2) based on the CpG clock assuming a ten-times higher transition rate [2,25] are approximately twice smaller (2.6, 1.8 and 0.9 Myr, respectively). These figures are consistent with separate amplification of Sb2 variants in these three lineages. In addition, the more recent retropositional activity of this subfamily (and thus shorter time period involved) could explain the lower number of Sb2 copies in apes.

DISCUSSION

The single base difference between the human and gorilla consensus sequences as well as the length variability that

Table 2. Divergence of *Alu* Sb2 repeats from the corresponding consensus sequences

	Human	Gorilla	Chimpanzee
substitutions at non-CpG positions	0.0075 (SD +/-0.0074)	0.0065 (SD +/-0.0067)	0.0038 (SD +/-0.0062)
transitions at CpG dinucleotides	0.0389 (SD +/-0.0373)	0.0250 (SD +/-0.0351)	0.0133 (SD +/-0.0186)

The numbers reported here are mean values obtained by averaging individual divergencies calculated for each sequence separately, such that the corresponding standard deviation (SD) can be derived. Almost the same values are obtained by dividing all substitutions in a given category by a total number of sequence positions considered.

distinguishes that of the chimpanzee indicate that the corresponding repeats amplified independently in these three primate lineages and should be considered as separate Alu subfamilies. To emphasize this fact we denote them by adding the abbreviated species name to that of the *Alu* subfamily, followed (optionally) by the information on sequence modification with respect to the standard consensus, here considered to be the human one. Thus, Sb2 subfamilies in *H.sapiens*, *G.gorilla* and *P.troglodytes* are denoted: Sb2*Hsa*, Sb2*Ggo*$-$(A176) and Sb2*Ptr*$-(\mathrm{A}133)_{0-8}$, or simply Sb2*Hsa*, Sb2*Ggo* and Sb2*Ptr*, respectively.

The estimated average age of the human repeats (Fig. 1 and Table 2) is consistent with the apparent absence of the orthologous inserts in apes (Fig. 2) in the few loci that were tested. The lack of insertional polymorphism in these loci suggests that Sb2 subfamily originated early in the history of the human lineage, while observations in the HD and CHEB loci indicate the persistence of their retropositional capacity. Gorilla Sb2 appears younger, and the chimpanzee Sb2 subfamily the youngest because of the lowest sequence divergence of its member repeats from the consensus. The different electrophoretic pattern in the common and pygmy chimpanzee agrees with the independent amplification of Sb2 repeats in these closely related species that separated about 3 Myr ago [26].

Estimations of the average age of Sb2 subfamilies in human, gorilla and chimpanzee lineages based on the CpG clock were systematically two times smaller than those calculated using a general clock. This difference could be explained by the uncertain ratio of CpG to the general (non-CpG) clock [2,25], and by small numbers that were considered in these calculations. However, this difference can be better accounted for by PCR and/or sequencing errors which result in an overestimation of the number of substitutions at non-CpG relative to CpG positions. In the latter, because of their inherently higher mutation rate, the contribution of false-positives will be lower, especially if only transitions $\mathrm{C} \rightarrow \mathrm{T}$ and $\mathrm{G} \rightarrow \mathrm{A}$ are considered. Obviously, this problem only arises in young repeats.

Does the length variability in the A-rich linker among chimpanzee repeats indicate that different versions of retroposing Sb2 RNA originated from different genomic locations? Because of the small size of Sb2*Ptr* subfamily it is more likely that there was possibly only one genomic 'pro-*Alu*' whose linker expanded over time. We do not know whether this expansion was due to chimpanzee-specific cellular factors or does it reflect the influence of the immediate genomic context. We believe that the comparison of Sb2 founder loci in the three primate lineages will shed light on this genomic instability and on the underlying molecular mechanism (work in progress). In addition, the data presented here indicate that Sb2 *Alu*'s are especially suitable for addressing questions concerning the expression and mobility of short retroelements. Apparently independent (or at least temporarily separated) activation of Sb2 retroposition in three primate lineages may help us to understand cis factors governing this process, once we isolate the genomic loci that transcribed retropositionally active Sb2 sequences in African great apes and human.

ACKNOWLEDGEMENTS

We are grateful to Drs Jurgen Brosius, Clement Lanthier and Robert Patenaude for their invaluable contribution by helping us to collect DNA samples from a variety of primate species. We would like to thank Guy Cardinal for sequencing the Sb2 clones. This work was supported by the grant from The Cancer Research Society, Inc.

REFERENCES

1. Schmid, C. and Maraia, R., (1992), Curr. Opin. Genet. Dev. 2, 874–882.
2. Labuda, D., Striker, G. (1989) Nucleic Acids Res. 17, 2477–2491.
3. Jurka, J., and Miloslajevic, A. (1991) J. Mol. Evol. 32, 105–121.
4. Labuda, D., Zietkiewicz, E. (1994) J. Mol. Evol. 39, 506–518.
5. Trabuchet, G., Chebloune, Y., Savatier, Y., Lachuer, L., Faure, J., Verdier, C. and Nigon, V.M. (1987) J. Mol. Evol. 25, 288–291.
6. Gibbs, P.E.M., Zielinski, R., Boyd, C. and Dugaiczyk, A. (1987) Biochem. 26, 1332– 1343.
7. Batzer, M.A., Kilroy, G.E., Richard, P.E., Shaikh, T.H., Deselle, T.D., Hoppens, C.L. and Deininger, P.L. (1990) Nucleic Acids Res 18, 6793–6798.
8. Matera, A.G., Hellman, U. and Schmid, C.W. (1990) Mol. Cell. Biol. 10, 5424–5432.
9. Batzer, M.A. and Deininger P.L. (1991) Genomics 9, 481–487.
10. Shen, M.R., Batzer, M.A. and Deininger, P.L. (1991) J. Mol. Evol. 33, 311–320.
11. Leeflang, E.P., Liu, W.–M., Chesnokov I.N. and Schmid, C.W. (1993) J. Mol. Evol. 37, 559–565.
12. Leeflang, E.P., Chesnokov I.N. and Schmid, C.W. (1993) J. Mol. Evol. 37, 566–572.
13. Batzer, M.A., Gudi, V.A., Mena J.C., Foltz D.W., Herrera R.J. and Deininger P.L (1991), Nucleic Acids Res. 19, 3619–3623.
14. Wallace, M.R., Andersen, L.B., Saulino, A.M., Gregory, P.E., Glover, T.W. and Collins, F.S. (1991) Nature 353, 864–866.
15. Muratani, L., Hada, K. Yamamoto, Y., Kaneko, T., Shigeto, Y., Ohue, T., Furuyama, J. and Higashino, K. (1991) Proc. Nat. Acad. Sci., U.S.A. 88, 11315–11319.
16. Vidaud, D., Vidaud, M., Bahnak, B.R., Siguret, V., Sanchez, S., Laurian, Y., Meyer, D., Goosens, M. and Lavergne, J.M. (1993) Eur. J. Hum . Genet., 1, 30–36.
17. Goldberg, Y.P., Rommens, J.M., Andrew,S.E., Hutchinson, G.B., Lin, B., Theilmann, J., Graham, R., Glaves, M.L., Starr, McDonald, E.H., Nasir, J., Schappert, K., Kalchman, M., Clarke, L. and Hayden, M.R. (1993) Nature 362, 370–373.
18. Hutchinson, G.B., Andrew, S.E., McDonald H., Goldberg, Y.P., Graham, R., Rommens J.M. and Hayden M.R. (1993), Nucleic Acids Res. 21, 3379–3383.
19. Jurka, J. (1993) Nucleic Acids Res. 21, 2252.
20. Zietkiewicz, E., Sinnett, D., Richer, C., Mitchell, G.A., Vanasse, M., Labuda, D., (1992) Human Genetics 89, 453–456.
21. Zietkiewicz, E., Labuda, M., Sinnett, D., Glorieux, F.H. and Labuda, D. (1992), Proc. Natl. Acad. Sci. USA, 89, 8448–8451.
22. Zietkiewicz, E.,; Makalowski, W., Mitchell. G. and Labuda, D. (1994) Science 265, 1110–1111.
23. Zietkiewicz, E., Akalin, N. and Labuda, D. (1994) Human Heredity 80, in press.
24. Britten, R.J. (1986) Science 231, 1393–1398.
25. Labuda, D., Sinnett, D., Richer, C., Deragon, J.–M. and Striker, G. (1991) J. Mol. Evol. 32, 405–414.
26. Sibley, C.G. and Ahlquist J.E. (1987) J. Mol. Evol. 26, 99–121.