Am. J. Hum. Genet. 73:1444–1451, 2003

Report

SVA Elements Are Nonautonomous Retrotransposons that Cause Disease in Humans

Eric M. Ostertag,[1,2] John L. Goodier,[1] Yue Zhang,[1] and Haig H. Kazazian Jr.[1]

[1]Department of Genetics, University of Pennsylvania School of Medicine, and [2]Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Philadelphia

L1 elements are the only active autonomous retrotransposons in the human genome. The nonautonomous *Alu* elements, as well as processed pseudogenes, are retrotransposed by the L1 retrotransposition proteins working in *trans*. Here, we describe another repetitive sequence in the human genome, the SVA element. Our analysis reveals that SVA elements are currently active in the human genome. SVA elements, like *Alus* and L1s, occasionally insert into genes and cause disease. Furthermore, SVA elements are probably mobilized in *trans* by active L1 elements.

The human genome is packed with repetitive DNA, predominantly the result of the activities of the LINE-1 (long interspersed nucleotide element–1) (or L1) retrotransposon. This retrotransposon is able to make an RNA copy of itself, then reverse transcribe the RNA into cDNA and integrate the cDNA copy into the genome. Two de novo insertions of L1 retrotransposons into the factor VIII gene of patients with hemophilia A demonstrated that they are a source of disease-causing mutation in humans (Kazazian et al. 1988; see Ostertag and Kazazian [2001*a*] for review). It was discovered subsequently that *Alu* elements also cause disease by inserting into genes (see Deininger and Batzer [1999] for review). Recent experimental data demonstrated that *Alu* elements are nonautonomous retrotransposons that use the L1 machinery for their own mobilization (Dewannieux et al. 2003). Therefore, in addition to causing disease by retrotransposition in *cis*, L1 elements are also responsible for disease by mobilization of *Alu* elements in *trans*. Mobilization of short interspersed nucleotide elements (SINEs) by LINEs has also been experimentally demonstrated in other organisms (Nikaido et al. 2002).

The SINE-R element was first reported as a novel retrotransposon derived from a human endogenous retrovirus, HERV-K10, and is present in the human genome at a frequency of 4,000–5,000 copies per haploid genome (Ono et al. 1987). Ono et al. (1987) noted that the elements ended in an A-rich sequence and were flanked by target-site duplications (TSDs). A second group described a SINE-R sequence within the human complement *C2* gene that was associated with a variable-number-of-tandem-repeats (VNTR) locus (Zhu et al. 1992). This sequence and another similar sequence from within the *HLA-RP1* (*STK19*) gene were subsequently analyzed by a third group, who reported that the SINE-R and VNTR sequences were also associated with *Alu*-like sequence and occasionally were associated with hexameric repeats (Shen et al. 1994). The entire structure was flanked by TSDs. Shen et al. (1994) used the term "SVA" (SINE-R, VNTR, and *Alu*) to describe this "composite retroposon."

Hassoun et al. (1994) reported the insertion of a novel mobile element into the α-spectrin gene (*SPTA1*) in a recent ancestor of a family with hereditary elliptocytosis (Hassoun et al. 1994). We have determined that this sequence was actually the result of an SVA-mediated transduction event. We have isolated and characterized the full-length SVA precursor that created the insertion SVA_{SPTA1}. We demonstrate that SVA insertions contain the hallmarks of retrotransposition in *trans* by the L1 element; that is, they are flanked by L1-like TSDs, they end in a poly A tail, they can transduce 3′ sequences, and they occasionally truncate and invert during insertion. Therefore, it is highly likely that SVA elements are retrotransposed by L1 proteins acting in *trans*. SVA elements join *Alu* elements as the only active nonautonomous retrotransposons that cause disease in humans.

A de novo insertion into the α-spectrin gene was caused by an SVA-mediated transduction.—We performed a

Received August 19, 2003; accepted for publication September 24, 2003; electronically published November 19, 2003.

Address for correspondence and reprints: Dr. H. H. Kazazian Jr., 475 Clinical Research Building, 415 Curie Boulevard, Philadelphia, PA 19104-6145. E-mail:kazazian@mail.med.upenn.edu

0002-9297/2003/7306-0021$15.00

BLAST search (Altschul et al. 1990) of GenBank, and we were able to find the likely precursor to a previously unknown sequence inserted into the α-spectrin gene of a family with hereditary elliptocytosis (fig. 1). The BLAST search returned a nearly perfect match to a sequence contained in contig AC016142. However, the sequence of the insertion was rearranged in relation to the sequence from the contig. The 5′ end of the insertion was inverted, and

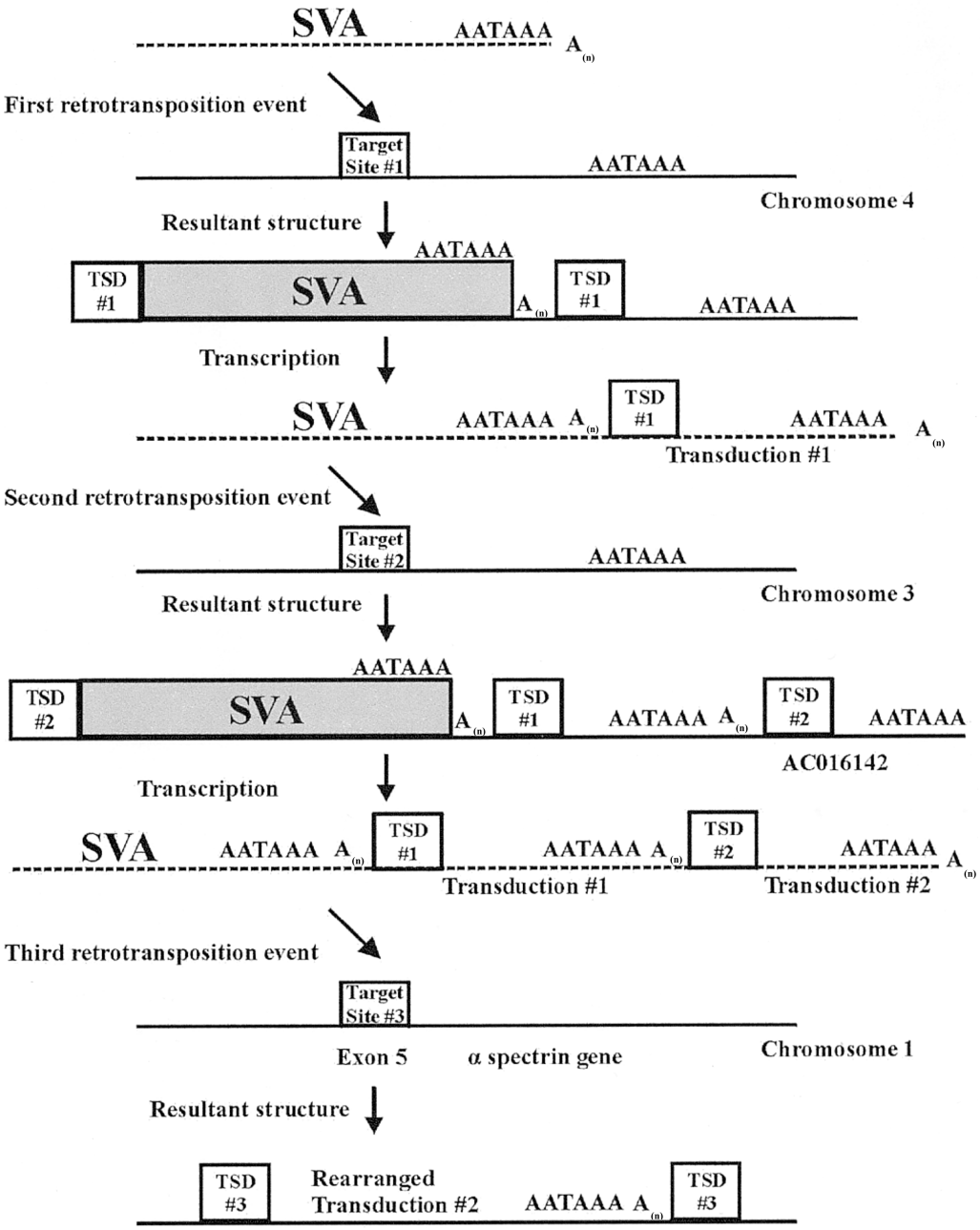

Figure 1 The sequence of SVA retrotransposition events. A full-length SVA element of unknown origin retrotransposed into the empty-site sequence on chromosome 4, represented on contigs AC037439 and AC068352. During subsequent transcription of the new SVA at this site, the SVA poly A was bypassed in favor of a downstream poly A, producing the first transduction event. The full-length SVA and transduction 1 inserted into a new genomic location at target site 2 on chromosome 3 to produce the sequence found in contig AC016142. During subsequent transcription of the new SVA on chromosome 3, both the SVA poly A and the poly A from transduction 1 were bypassed in favor of a poly A farther downstream, producing the second transduction event. The transcript containing a full-length SVA element and both transduction events inserted into target site 3 on chromosome 1, which is in exon 5 of the α-spectrin gene. However, the integration process resulted in a structure that was 5′ truncated and inverted compared with the precursor, a common process in L1-mediated retrotransposition. RNA is represented by a dashed line; DNA is represented by a solid line.

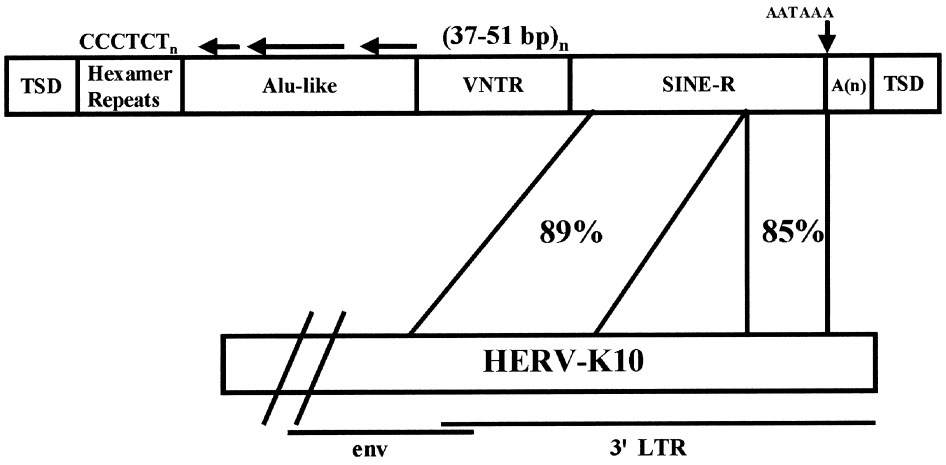

Figure 2 Structure of a full-length SVA element. Most SVA elements in the genome are flanked by L1-like TSDs. The 5′ end of a full-length SVA element consists of a variable number of (CCCTCT) hexameric repeats, followed by sequence with homology to antisense *Alu* sequence (arrows represent approximate areas of homology), a VNTR region, a SINE-R region, and an AATAAA poly A signal, followed immediately by a poly A tail, that is A(n). SINE-R sequence is related to the 3′ LTR sequence of HERV-K10 (regions of identity and percent identity indicated).

there was a 22-bp deletion of the sequence at the point of inversion. This type of rearrangement is common during L1 retrotransposition (Ostertag and Kazazian 2001*b*), suggesting that perhaps the inserted sequence was the result of an L1-mediated transduction (Moran et al. 1999; Goodier et al. 2000; Pickeral et al. 2000) that had inverted during the insertion process. If this hypothesis were correct, then one would predict that an active, full-length L1 element would be located 5′ from the transduced sequence. However, when we examined the 5′ sequence, we did not find a full-length L1 element; rather, we found a full-length SVA element, which suggests that the insertion into the α-spectrin gene was the result of an SVA-mediated retrotransposition event.

The full-length SVA element in contig AC016142 contained evidence of a previous transduction event. In other words, the sequence that was transduced during the retrotransposition event into the α-spectrin gene was a secondary transduction event. Such composite transductions occasionally occur during L1 retrotransposition (Goodier et al. 2000). The evidence of the original transduction was the presence of 183 bp of sequence, which we will refer to as "transduction 1," that lay between the full-length SVA element and the secondary transduction event of 599 bp, which we will refer to as "transduction 2." A transduced sequence acts as a molecular address that permits the identification of the precursor to the transduction event. Therefore, we performed a second BLAST search and identified the location of the first transduction. The transduction 1 sequence was an identical match to sequence from contigs AC037439 and AC068352. However, neither of these contigs contained a full-length SVA element; rather, they represented the "empty site" of this insertion (fig. 1).

Structure of SVA elements.—We acquired BAC clone RP11-166N6 (accession number AC016142) (Research Genetics), isolated the full-length SVA precursor that created SVA_{SPTA1}, and cloned it into pBluescript (Stratagene) with standard molecular subcloning techniques to create pBS-SVA. We sequenced pBS-SVA using the following oligonucleotide primers: SVA5′Flank(F) (5′-CAT GGA TGG TGT CAA GCT AC-3′), SVA(*Alu*)R (5′-CAC CAC TGA GCA CTG AGT GAA C-3′), and T7 (5′-GTA ATA CGA CTC ACT ATA GGG C-3′).

After we cloned and restriction mapped the precursor to the α-spectrin insertion, it became clear that there were errors in the assembly of contig AC016142. The GenBank sequence for contig AC016142 apparently had incorrect 5′ flanking sequence. However, this error has been corrected. From the correct sequence, we were able to surmise the structural features of a full-length SVA element (fig. 2). At their 5′ ends, full-length SVA elements have hexameric (CCCTCT) repeats. This region is followed by (1) a region containing an antisense *Alu* sequence, (2) a VNTR region containing multiple copies of a 35–50-bp repeat, (3) a SINE-R sequence, and (4) a polyadenylation signal and a poly A tail.

The exact number of hexameric repeats present at the 5′ end of SVA_{SPTA1} has been difficult to determine because the repeat structure is intractable to sequencing in either direction. However, from the sequencing results, we know that there are a minimum of 10 repeats. In addition, we performed a PCR across the hexamers, using primers SVA5′Flank(F) and SVA(*Alu*)R. From the size of this PCR product, we estimate that there are ⩽13 hexameric repeats (data not shown).

SVA_{SPTA1} contains 35 repeats in the VNTR region. We aligned the repeats, using ClustalW (fig. 3*A*). There are

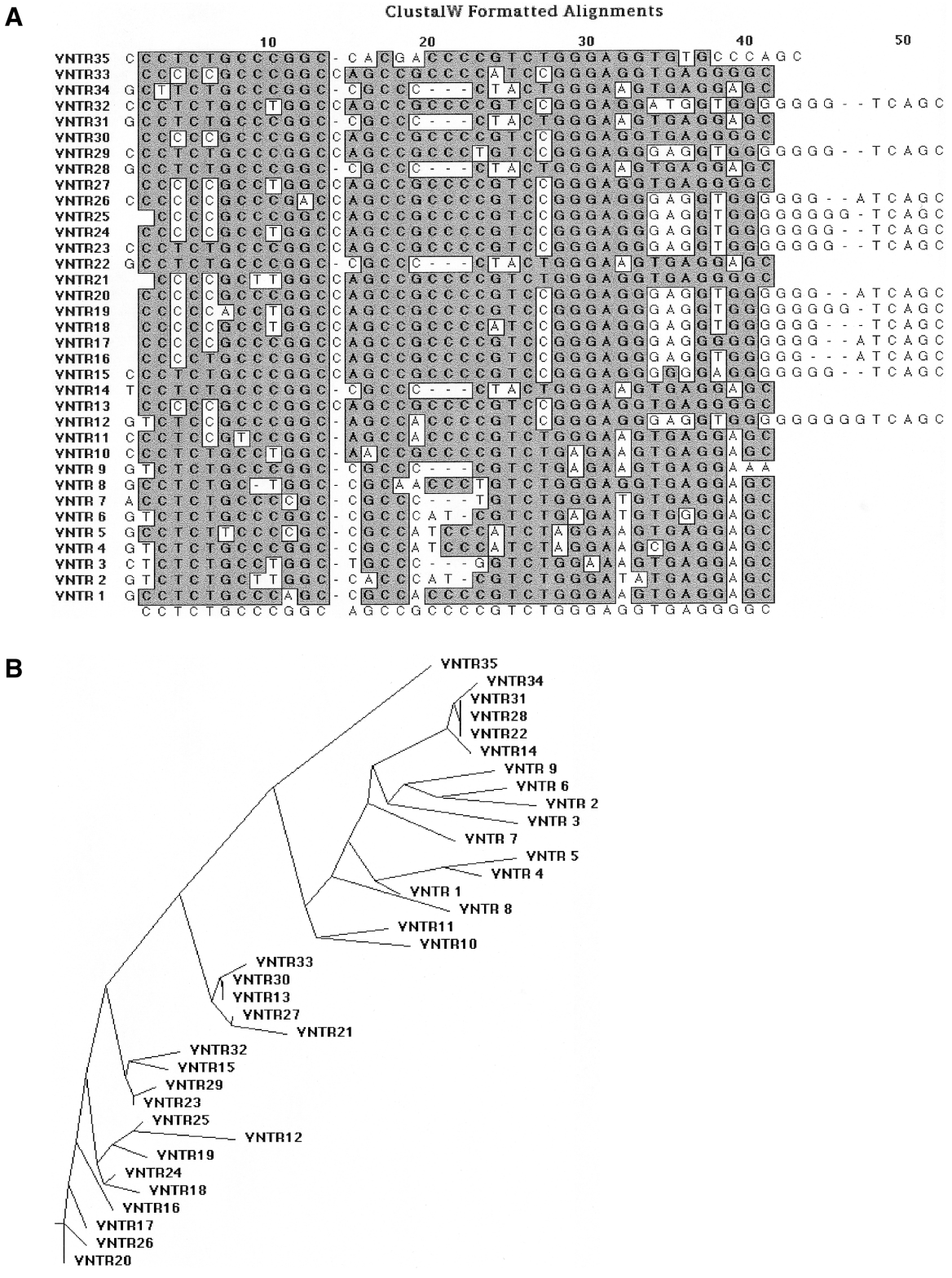

Figure 3 Analysis of the SVA VNTRs. ClustalW alignment of the 35 SVA_{SPTA1} VNTRs (*A*) and associated cladogram (*B*), which shows evidence of VNTR expansion by unequal homologous recombination. Notice that the individual VNTRs from a block of VNTRs (18–23) are most closely related to the corresponding individual VNTRs from block 24–29; that is, VNTR18 is most similar to VNTR24, VNTR19 is most similar to VNTR25, VNTR20 is most similar to VNTR26, etc.

two general types of VNTR, one with a range of 36–42 bp and the other with a range of 49–51 bp. We performed a cladogram on the VNTRs, using MacVector software (fig. 3*B*). The relationship of the VNTRs suggests that the region has expanded as the result of mispairing of the VNTR regions, followed by unequal homologous recombination. For example, the individual VNTRs from a block of six VNTRs (18–23) are most closely related to the corresponding individual VNTRs from block 24–29. One would predict that the unequal homologous recombination event that produced the duplication of the six-VNTR–block region also produced a deletion of the six VNTRs in the other SVA allele. Unequal homologous recombination has been proposed elsewhere to describe features of simple repetitive DNA elements (Levinson and Gutman 1987). Perhaps SVA elements are hotspots for unequal homologous recombination in the human genome.

Copy number and similarity.—We estimated the number of full-length and truncated SVA elements in the genome by using variable amounts of the 5′ end of the element (including as much as the first 380 bp of the *Alu*-like region) or the 3′ end of the element (as much as the last 380 bp of the SINE-R region) to perform BLAST searches of the draft human genome sequence in GenBank. Our BLAST searches by use of sequences from the 3′ end of SVA_{SPTA1} suggest that there are ~3,500–7,500 SVA sequences in the diploid human genome, which is in agreement with the estimate of the number of SINE-R sequences performed by Ono et al. (1987) with filter hybridization screens (Ono et al. 1987). Estimates through use of the 5′ end of the element suggest that >2,000 of the SVA elements are full length. BLAT searches (Sugnet et al. 2002) simultaneously using both ends of the element confirmed the presence of at least several hundred full-length elements. We found many copies of SVA elements with >99% identity to recent SVA insertions in the SINE-R region, and all SVA elements were >89% identical to each other.

SVA retrotransposons have evolved recently.—Retrotransposons that have been in the genome for a long time accumulate random mutations. Therefore, sequence divergence within a family of retrotransposons is correlated with the age of the retrotransposon (Boissinot et al. 2000). The lack of any SVA elements with high sequence divergence probably indicates that SVAs have evolved recently. On the other hand, the low nucleotide divergence may be the result of gene conversion, which has been documented in other human retroelements (Myers et al. 2002; Roy-Engel et al. 2002*a*). Stronger evidence for their young age comes from phylogenetic studies of the SINE-R component and its presumptive precursor, HERV-K10 (Ono et al. 1987; Zhu et al. 1992). HERV-K10 is a retroviral-like sequence present in the genomes of humans, chimpanzees, gorillas, orangutans, and at least some Old World monkeys (Ono et al. 1986; Zhu et al. 1994). One study of SINE-R sequences by Southern analysis found them present only in humans, chimpanzees, and gorillas (Zhu et al. 1994). However, a subsequent PCR analysis of SINE-R sequences found SINE-R in all hominoid primates (Kim et al. 1999). Since SVA elements are all >89% similar in sequence to each other and are present only in hominoid primates, these elements are quite young by evolutionary standards, probably <15 million years old. The very young evolutionary age of the SVA element may present a unique opportunity to study the entire evolutionary history of a human retrotransposon. Such an analysis could not be performed on the L1 retrotransposon, because L1 has been in the genome for at least several hundred million years, and copies >200 million years have diverged to the point that they are now unrecognizable (Smit 1999). In addition, these elements may be valuable as markers for primate or human phylogenetic and population studies, as has been the case for the *Alu* element (Bamshad et al. 2003; Watkins et al. 2003).

SVA elements are active retrotransposons.—The several thousand copies of SVA in the human genome are a relatively small number when compared with the 500,000 copies of L1 and 1,100,000 copies of *Alu* (Lander et al. 2001). There are ~170-fold more L1 sequences and >300-fold more *Alu* sequences in the genome. However, this may reflect an evolutionarily young age of the SVA element rather than a lack of activity. There are several reasons to believe that SVA elements are quite active in the genome. One of the few ways to estimate retrotransposition frequency is by the number of recent or de novo disease-causing mutations that the element has caused by its retrotransposition. A recent survey of retrotransposon insertions reported 13 cases of recent or de novo disease-causing L1 insertions and 18 cases of *Alu* insertion (Ostertag and Kazazian 2001*a*). A few additional insertions have been described in the past 2 years. In addition to the insertion into the α-spectrin gene, the SVA element has caused at least two other disease-causing insertions. A de novo SVA insertion into the BTK gene of a patient with X-linked agammaglobulinemia was reported elsewhere as a SINE-R insertion (Rohrer et al. 1999). In addition, an SVA insertion into the *fukutin* gene is the cause of Fukuyama-type congenital muscular dystrophy in a large group of Japanese patients (Kobayashi et al. 1998). The insertion into the *fukutin* gene was not a de novo event, but it is estimated to have occurred ~102 generations ago (Colombo et al. 2000). Therefore, despite its much lower copy number than those of L1 and *Alu* elements, SVA has an estimated cumulative retrotransposition activity that is only fivefold less than L1 and sevenfold less than *Alu*.

The SVA elements in the genome have high sequence identity to each other, which suggests that they have

retrotransposed recently. In addition, because the insertion in the α-spectrin gene originated from a transcript with a composite transduction, we have been able to find not only the precursor to the insertion but also the empty-site locus of the precursor to the precursor. The fact that we found the empty site suggests that the SVA element at that locus is polymorphic in the population with regard to its presence or absence. Therefore, we have evidence of multiple rounds of retrotransposition in the recent past. Other young, active retrotransposons—such as Ta-0 and Ta-1, the youngest human L1 subfamilies—also display both low sequence divergence and a high degree of polymorphism (Boissinot et al. 2000; Sheen et al. 2000; Myers et al. 2002; Ovchinnikov et al. 2002; Brouha et al. 2003). A similar situation is seen among *Alu* elements (reviewed in Batzer and Deininger [2002]).

SVA elements are nonautonomous and bear all the hallmarks of mobilization in trans *by the L1 retrotransposition machinery.*—Many characteristics of SVA insertions are reminiscent of L1 insertions. First, some insertions are 5′ truncated (e.g., the BTK insertion), and others are both 5′ truncated and inverted (e.g., the α-spectrin insertion). These are both common processes during L1 retrotransposition. Second, they end in a poly A tail directly following a poly A signal. This may indicate that the adenine bases between the poly A signal and the cleavage site have been positively selected for by L1 target-primed reverse transcription (TPRT), the process that L1 elements are thought to use when integrating into the genome (Luan et al. 1993). Third, they occasionally bypass their own poly A signal in favor of a downstream signal, which results in a transduction event. Therefore, their retrotransposition does not depend on any specific 3′ sequence. The lack of any 3′ sequence requirements, other than a poly A tail, for retrotransposition is consistent with L1 TPRT. Last, they are flanked by 10-20-bp long TSDs, and the sequences of predicted cleavage sites are the same as those of the L1 endonuclease, which is 3′-AA/TTTT-5′ and minor variations (Feng et al. 1996; Jurka et al. 1997). For example, the predicted cleavage site of the α-spectrin insertion is 3′-TA/TTCT-5′, and that of the SVA_{SPTA1} is 3′-AA/TTTT-5′. These data are a strong indication that the L1 retrotransposition machinery mediates their retrotransposition.

Minimum sequence requirements of a human nonautonomous retrotransposon.—A nonautonomous retrotransposon depends on the retrotransposition machinery of an autonomous retrotransposon to propagate itself. L1 is the only active retrotransposon in the human genome. Therefore, any human nonautonomous element must contain the sequence requirements that would allow it to be retrotransposed by L1. One obvious requirement is an internal promoter that permits expression in the same tissue and at the same time as L1 retrotransposition is occurring. The L1 element has a germline-specific promoter and is able to retrotranspose in the germline (Ostertag et al. 2002). Presumably, the SVA promoter also expresses in the human germline. A nonautonomous element must also have any sequences required for transcription located downstream of the transcription initiation site; otherwise, they would be lost on retrotransposition. Full-length SVA elements in the genome tend to begin within the hexameric region, suggesting that transcription is directed to initiate within this region. Further, a human nonautonomous retrotransposon must have the ability to be reinserted into the genome by the L1 TPRT mechanism. Apparently, the only absolute sequence requirement for L1-mediated TPRT is the presence of a poly A tail (Roy-Engel et al. 2002*b;* Dewannieux et al. 2003). SVA elements contain a poly A tail that is very similar to that of L1.

There may be an additional requirement of a human nonautonomous retrotransposon. The only nonautonomous retrotransposons in the genome with evidence of current activity are the *Alu* elements and the SVA elements. It would seem that there is something special about *Alu* transcripts and SVA transcripts that allows them to retrotranspose at such high frequencies, when almost all other transcripts are not *trans* complemented at appreciable levels (Wei et al. 2001), including those that are more highly expressed in the germline. It is interesting that *Alus* and SVAs both have *Alu* sequence components. Perhaps the *Alu* sequences are important in the *trans* complementation by L1, placing the element RNA in close proximity to the L1 RNA, either on the ribosome or within a ribonucleoprotein particle (Boeke 1997).

Acknowledgments

H.H.K. was supported by a National Institutes of Health grant. E.M.O. was supported by a Howard Hughes Medical Institute predoctoral fellowship.

Electronic-Database Information

Accession numbers and the URL for data presented herein are as follows:

GenBank, http://www.ncbi.nih.gov/Genbank/ (for contigs AC037439 and AC068352 and BAC clone RP11-166N6 [contig AC016142])

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bamshad MJ, Wooding S, Watkins WS, Ostler CT, Batzer MA, Jorde LB (2003) Human population genetic structure and

inference of group membership. Am J Hum Genet 72:578–589

Batzer MA, Deininger PL (2002) Alu repeats and human genomic diversity. Nat Rev Genet 3:370–379

Boeke JD (1997) LINEs and Alus—the polyA connection. Nat Genet 16:6–7

Boissinot S, Chevret P, Furano AV (2000) L1 (LINE-1) retrotransposon evolution and amplification in recent human history. Mol Biol Evol 17:915–928

Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH Jr (2003) Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci USA 100:5280–5285

Colombo R, Bignamini AA, Carobene A, Sasaki J, Tachikawa M, Kobayashi K, Toda T (2000) Age and origin of the FCMD 3′-untranslated-region retrotransposal insertion mutation causing Fukuyama-type congenital muscular dystrophy in the Japanese population. Hum Genet 107:559–567

Deininger PL, Batzer MA (1999) Alu repeats and human disease. Mol Genet Metab 67:183–193

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

Feng Q, Moran JV, Kazazian HH Jr, Boeke JD (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905–916

Goodier JL, Ostertag EM, Kazazian HH Jr (2000) Transduction of 3′-flanking sequences is common in L1 retrotransposition. Hum Mol Genet 9:653–657

Hassoun H, Coetzer TL, Vassiliadis JN, Sahr KE, Maalouf GJ, Saad ST, Catanzariti L, Palek J (1994) A novel mobile element inserted in the alpha spectrin gene: spectrin dayton. A truncated α spectrin associated with hereditary elliptocytosis. J Clin Invest 94:643–648

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

Kazazian HH Jr, Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE (1988) Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332:164–166

Kim HS, Takenaka O, Crow TJ (1999) Cloning and nucleotide sequence of retroposons specific to hominoid primates derived from an endogenous retrovirus (HERV-K). AIDS Res Hum Retroviruses 15:595–601

Kobayashi K, Nakahori Y, Miyake M, Matsumura K, Kondo-Iida E, Nomura Y, Segawa M, Yoshioka M, Saito K, Osawa M, Hamano K, Sakakihara Y, Nonaka I, Nakagome Y, Kanazawa I, Nakamura Y, Tokunaga K, Toda T (1998) An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. Nature 394:388–392

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72:595–605

Moran JV, DeBerardinis RJ, Kazazian HH Jr (1999) Exon shuffling by L1 retrotransposition. Science 283:1530–1534

Myers JS, Vincent BJ, Udall H, Watkins WS, Morrish TA, Kilroy GE, Swergold GD, Henke J, Henke L, Moran JV, Jorde LB, Batzer MA (2002) A comprehensive analysis of recently integrated human Ta L1 elements. Am J Hum Genet 71:312–326

Nikaido M, Harada M, Matsumura S, Lin LK, Wu Y, Hasegawa M, Okada N, Kajikawa M (2002) LINEs mobilize SINEs in the eel through a shared 3′ sequence. J Mol Evol 55:284–301

Ono M, Kawakami M, Takezawa T (1987) A novel human nonviral retroposon derived from an endogenous retrovirus. Nucleic Acids Res 15:8725–8737

Ono M, Yasunaga T, Miyata T, Ushikubo H (1986) Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. J Virol 60: 589–598

Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, Kazazian HH Jr (2002) A mouse model of human L1 retrotransposition. Nat Genet 32:655–660

Ostertag EM, Kazazian HH Jr (2001*a*) Biology of mammalian L1 retrotransposons. Ann Rev Genet 35:501–538

———(2001*b*) Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. Genome Res 11:2059–2065

Ovchinnikov I, Rubin A, Swergold GD (2002) Tracing the LINEs of human evolution. Proc Natl Acad Sci USA 99: 10522–10527

Pickeral OK, Makalowski W, Boguski MS, Boeke JD (2000) Frequent genomic DNA transduction driven by LINE-1 retrotransposition. Genome Res 10:411–415

Rohrer J, Minegishi Y, Richter D, Eguiguren J, Conley ME (1999) Unusual mutations in Btk: an insertion, a duplication, an inversion, and four large deletions. Clin Immunol 90:28–37

Roy-Engel AM, Imataka H, Sonenberg N, Deininger P (2002*a*) Non-traditional Alu evolution and primate genomic diversity. J Mol Biol 321:423–432

Roy-Engel AM, Salem A-H, Oyeniran OO, Deininger L, Hedges DJ, Kilroy GE, Batzer MA, Deininger PL (2002*b*) Active Alu element "A-tails": size does matter. Genome Res 12:1333–1344

Sheen FM, Sherry ST, Risch GM, Robichaux M, Nasidze I, Stoneking M, Batzer MA, Swergold GD (2000) Reading between the LINEs: human genomic variation induced by LINE-1 retrotransposition. Genome Res 10:1496–1508

Shen L, Wu LC, Sanlioglu S, Chen R, Mendoza AR, Dangel AW, Carroll MC, Zipf WB, Yu CY (1994) Structure and genetics of the partially duplicated gene RP located immediately upstream of the complement C4A and the C4B genes in the HLA class III region: molecular cloning, exon-intron structure, composite retroposon, and breakpoint of gene duplication. J Biol Chem 269:8466–8476

Smit AF (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr Opin Genet Dev 9:657–663

Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D, Kent WJ (2002) BLAT—the BLAST-like alignment tool. Genome Res 12:996–1006

Watkins WS, Rogers AR, Ostler CT, Wooding S, Bamshad MJ, Brassington A-ME, Carroll ML, Nguyen SV, Walker JA, Prasad BVR, Reddy PG, Das PK, Batzer MA, Jorde LB (2003) Genetic variation among world populations: inferences from 100 *Alu* insertion polymorphisms. Genome Res 13:1607–1618

Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JD, Moran JV (2001) Human L1 retrotransposition: *cis*-preference vs. *trans*-complementation. Mol Cell Biol 21:1429–1439

Zhu ZB, Hsieh SL, Bentley DR, Campbell RD, Volanakis JE (1992) A variable number of tandem repeats locus within the human complement C2 gene is associated with a retroposon derived from a human endogenous retrovirus. J Exp Med 175:1783–1787

Zhu ZB, Jian B, Volanakis JE (1994) Ancestry of SINE-R.C2 a human-specific retroposon. Hum Genet 93:545–551