

.....

Retrotransposable Elements and Human Disease

P.A. Callinan, M.A. Batzer

Department of Biological Sciences, Biological Computation and Visualization
Center, Center for BioModular Multi-Scale Systems, Louisiana State
University, Baton Rouge, La., USA

Abstract

Nearly 50% of the human genome is composed of fossils from the remains of past transposable element duplication. Mobilization continues in the genomes of extant humans but is now restricted to retrotransposons, a class of mobile elements that move via a copy and paste mechanism. Currently active retrotransposable elements include Long INterspersed Elements (LINEs), Short INterspersed Elements (SINEs) and SVA (SINE/VNTR/*Alu*) elements. Retrotransposons are responsible for creating genetic variation and on occasion, disease-causing mutations, within the human genome. Approximately 0.27% of all human disease mutations are attributable to retrotransposable elements. Different mechanisms of genome alteration created by retrotransposable elements include insertional mutagenesis, recombination, retrotransposition-mediated and gene conversion-mediated deletion, and 3' transduction. Although researchers in the field of human genetics have discovered many mutational mechanisms for retrotransposable elements, their contribution to genetic variation within humans is still being resolved.

Copyright © 2006 S. Karger AG, Basel

Transposable Elements in the Human Genome

Almost the entire human genome is ubiquitously littered with the skeletons of mobile elements, which all told, account for a staggering 45% of the sequence content [1]. Mobile elements successfully accumulated in genomes during eukaryotic evolution and are grouped into one of two different classes: DNA transposons or retrotransposons.

DNA transposons constitute 3% of the human genome [1] and although they are represented by inactive fossils in humans, DNA transposons remain

active in the genomes of plants, flies and bacteria [2–4]. Retrotransposons, on the other hand, are currently actively mobilizing within the human genome and comprise approximately 40% of the DNA sequence [1]. Due to the current propagation of retrotransposons in humans, they will be the focus of this review.

Retrotransposons, by definition, mobilize via an RNA intermediate that is subsequently reverse transcribed into a cDNA copy using a mechanism termed Target Primed Reverse Transcription (TPRT) [5]. This copy and paste mechanism of mobilization results in the spread of retrotransposons to new genomic locations. Retrotransposable elements are categorized based on their ability to mobilize. Long INterspersed Elements (LINEs) are autonomous retrotransposons that encode the enzymatic machinery required for their propagation [6]. Short INterspersed Elements (SINEs), such as *Alu*, and SVA (SINE/VNTR/*Alu*) elements, are non-autonomous and thus require the enzymatic machinery of LINE elements for retrotransposition [7, 8].

Over the last quarter century, many ideas concerning the function of mobile elements have been put forth. Orgel and Crick were proponents of the idea that mobile elements served no function and resided as parasitic entities within the genome, without contributing to the evolutionary well-being of the organism [9]. Others have hypothesized that mobile elements function as origins of replication [10], chromosomal band-aids [11] and mediators of translational activation [12].

Despite disagreement over the function of mobile elements, they constitute an interesting source of human genomic variation and occasionally, disease. Here we present an overview of the contribution of mobile elements, in particular, retrotransposable elements, to genetic disease in *Homo sapiens*.

Autonomous Retrotransposons and Disease

Long INterspersed Elements

Computational analyses of the human genome have shown that L1 elements have reached a copy number in excess of 500,000 and comprise some 17% of the genomic sequence [1]. Numerous studies indicate that some subclasses of L1 element are still actively expanding by retrotransposition in extant human genomes [6]. Retrotranspositionally active L1 elements are approximately 6 kb in length, as shown in figure 1a. Evidence suggests that L1 elements have orchestrated large-scale alterations in the genomic architecture of human beings, as they are the major source of reverse transcriptase, upon which other retrotransposable elements and processed pseudogenes have amplified [6]. As a result, L1 elements are both directly and indirectly responsible for the vast

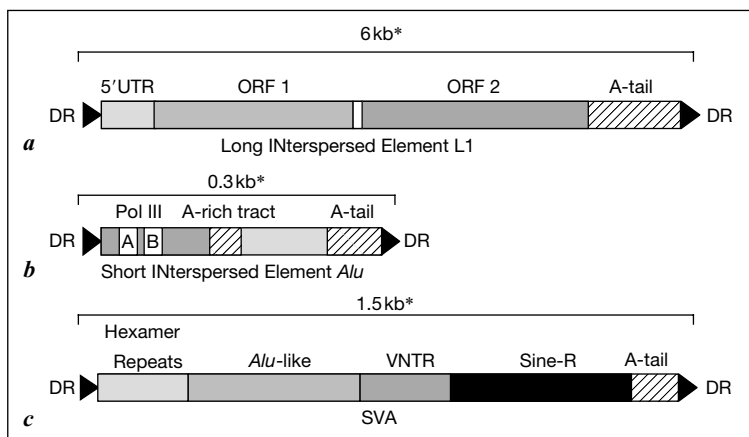


Fig. 1. Active retrotransposons within the human genome. **a** Long Interspersed Element L1. L1s are approximately 6 kb long and possess a 5' UTR, within which resides an RNA polymerase II promoter. Full-length elements encode two open reading frames that produce a reverse transcriptase and endonuclease, as well as an RNA binding protein. Each L1 element has a 3' UTR, an oligo-dA tail and is flanked by direct repeat sequences (DR). **b** Short Interspersed Element *Alu*. *Alu* SINEs are approximately 300 bp long and comprise two arms separated by a middle A-rich tract. They possess an RNA polymerase III (POL III) promoter (A and B box), in addition to a variable length oligo-dA rich tail. *Alu* elements are flanked by short direct repeats (DR). **c** SVA. Full-length SVA elements are approximately 1.5 kb long, and are composed of several repeat elements: a CCCTCT_n hexamer repeat, an anti-sense *Alu*, a variable number of tandem repeats, and a SINE-R element. SVA elements possess an oligo dA-rich tail and are flanked by short direct repeats (DR). * Not drawn to scale.

majority of retrotransposable element-derived variation and disease within the human genome. The propagation of L1 has resulted in disease-causing de novo insertions within genes, many of which disrupt exons or alter RNA splicing in the mutant alleles. In addition, the 500,000 L1 elements in the human genome provide long regions of sequence identity that represent numerous sites for unequal homologous recombination and mutation. Despite their vast numbers and retrotransposition activity, L1 elements are directly responsible for less than 20% of all retrotransposable element-related human diseases, even though experimental evidence suggests that L1s demonstrate a *cis* preference for their own replication machinery (see review [6]). The paucity of disease-causing L1 insertions may stem from L1 AT-rich insertion preference, essentially sidestepping the sensitive coding regions of the genome, or perhaps new L1 insertions are subject to appreciable amounts of negative selection because of their size. Additionally, distant L1 spacing may mean that recombination between L1

elements would induce fatal genetic damage and be eliminated. Due to the paucity of disease-causing L1 recombination events, we will not cover this particular mechanism here. Instead, we will focus on what is currently known concerning L1 retrotransposition, retrotransposition-mediated genomic deletion and 3' transduction and their contribution to human diseases.

L1 Retrotransposition

Newly inserted L1 elements have induced disease in sixteen separate documented cases and the vast majority of these elements belong to one of the youngest L1 subfamilies, termed Ta. The L1 Ta subfamily is approximately 2 million years old and shows a high level of polymorphism (insertion presence/absence) in diverse human populations [13].

In 2001, a comprehensive study of newly inserted L1 elements and related diseases was published [6]. The data gathered in this study indicated that nine out of the thirteen disease-causing L1 insertions discovered until that time disrupted sex-linked genes, namely Factor VIII, dystrophin or *CYBB* [6]. This observation suggests that some genes are hotspots for mobile element integration, or that the ensuing genic damage was easily detected due to their genomic position on the X chromosome, i.e. through ascertainment bias. Since the review in 2001 [6], three new cases of L1 induced X-linked genetic disease have been discovered. The first case describes an L1 insertion into the *RPS6KA3* gene causing Coffin-Lowry syndrome [14]. Second, a disruption of intronic splicing through an L1 insertion into the *CHM* gene causing choroideremia [15], and finally, a case of hemophilia B induced by L1 disruption of the Factor IX gene [16].

L1 disease-causing insertions have been mapped to both the exons and introns of genes. Most exonic L1 integrations are presumably lethal due to the introduction of premature stop codons and are likely eliminated from the population. However, nine instances of exonic integration have resulted in phenotypically tolerable diseases in humans. Intronic L1 insertions may also be lethal, but some studies have documented the existence of tolerable intronic insertions [6]. L1 elements have recently been shown to reduce mRNA transcript levels due to their presence within introns [17]. This phenomenon is related to the inefficiency of RNA polymerase II to transcribe through L1 elements [17]. Researchers suggest that L1 elements may act as 'molecular rheostats' by directly altering gene expression in this way [17]. Another study also recently demonstrated that RNA polymerase II transcription of L1 elements is adversely affected due to multiple termination and polyadenylation signals along the length of the L1 element [18]. It was proposed that premature RNA polymerase II termination could be a way that L1 elements limit their damage to host genomes [18]. At the same time, it would also mean that the stalling of polymerase molecules along L1 sequence would increase the negative impact of L1

insertions into genes [18]. Intergenic insertions of L1 may also alter gene expression throughout the human genome. L1 elements possess one RNA polymerase II promoter on their sense strand and another on their anti-sense strand that have been implicated in the enhancement of some genes (Factor IX and apolipoprotein A genes) and in the formation of chimeric mRNA transcripts [6]. Given the high insertion polymorphism levels of young L1 elements within the human genome, intronic and intergenic insertions could profoundly influence gene expression on both the individual and population level.

L1 Retrotransposition-Mediated Deletion

L1 retrotransposition-mediated deletion was first reported in 2002, where L1 integrations within cultured human cells resulted in target site deletions spanning from 1 bp to 70,000 bp at a rate of about 10% [19–21]. These studies hinted at the vast impact that L1 retrotransposition-mediated deletion may have had on primate genomes. If 10% of the L1 retrotranspositions induced deletions, then over 5,000 L1 retrotranspositions would be responsible for eliminating megabases of primate genomic DNA.

Retrotransposition events that resulted in deleted target site DNA were found to possess atypical characteristics, including a lack of target site duplications (TSDs), non-canonical L1 EN (endonuclease) nick sites and sometimes the absence of an oligo-dA rich tail (see [11, 20, 21]). Researchers proposed two models, based on evidence from in vitro retrotransposition studies, to help explain the mechanism for the insertion-deletion events. The first model proposed that L1 EN nicking variation on the top strand could account for TSD-less L1 element structure, in addition to genomic deletion at the site of insertion [11, 20, 21]. The second mechanism suggested that L1 reverse transcriptase could initiate TPRT from existing breaks in the genome, not depending on L1 EN for the initial nick [11]. Recently, a third model was formulated to explain the mechanism of retrotransposition-mediated deletion, named promiscuous TPRT (pTPRT) [22]. This model states that a retrotransposable element RNA transcript may hybridize to a region of genomic DNA downstream of a genomic break in order to initiate TPRT. The displaced single stranded DNA is removed through enzymatic degradation or by mechanical force, in order to create the target site deletion.

A recent survey of L1 disease-causing insertions reported two instances of retrotransposition-mediated deletion in humans: a 1-bp deletion in the *DMD* gene and another 6-bp deletion in the *FCMD* gene that resulted in Duchenne muscular dystrophy and Fukuyama-type congenital muscular dystrophy, respectively [23, 24]. In both cases, the disease phenotype resulted from the L1 element insertion, rather than through deletion of genomic sequence at the target site. These two cases are among only six other published in vivo examples

of L1 retrotransposition-mediated deletion in the human genome to date [25, 26]. Further research is underway at this time to determine the frequency of L1 retrotransposition-mediated deletion in the native human genome and its resultant impact on genomic instability and evolution.

L1-Mediated 3' Transduction

A decade ago, a mechanism was detected by which L1 alters the primate genome. It was termed 3' transduction [27]. The discovery of 3' transduction coincided with the insertion of L1 into the dystrophin gene, manifesting muscular dystrophy in a single human individual [27]. Since then, cell based studies have documented the ability of L1 elements to shuffle genomic DNA, including exons, using this mechanism (see [28]). During 3' transduction, a read-through transcript of the L1 element transcribes flanking genomic material downstream by virtue of a weak L1 termination and polyadenylation signal. Transduction of adjacent genomic DNA by L1 elements may result in the creation of new exons and in the alteration of gene expression through promoter and enhancer shuffling. Computational analyses have indicated that L1-mediated transduction of genomic material may occur at a rate of one in every five L1 retrotransposition events and that approximately 1% of the human haploid genome may have arisen by this mechanism [29]. In some instances, due to the severe truncation of L1 elements upon reverse transcription, it is possible that the transduced sequence will not reside adjacent to its L1 element thereby artificially reducing estimates of the impact that 3' transduction has had on the architecture of the human genome.

Non-Autonomous Retrotransposons and Disease

Alu Elements

The *Alu* family represents an enormously successful lineage of retrotransposons, whose origin and amplification coincided with the radiation of primates some 65 million years ago [5]. *Alu* elements are non-autonomous retrotransposons that mobilize in a copy and paste fashion. They are approximately 300 bp long and comprise two nearly identical arms separated by a middle A-rich tract, in addition to a 3' oligo dA-rich tail (fig. 1b). Recent data suggest that only a fraction of *Alu* elements, termed source genes, are retrotranspositionally competent and responsible for producing over one million *Alu* copies within the primate order [5]. Although the exact characteristics of a source gene are unclear, *Alu* element age, RNA polymerase III promoter integrity and the length and homogeneity of the oligo-dA rich tail are considered major factors influencing retrotransposition potential [5]. *Alu* elements have continued to mobilize throughout the evolution of primates, as evidenced by human lineage-specific elements.

These elements are absent from orthologous loci in non-human primates and exhibit high levels of polymorphism with respect to their insertion presence and absence in different human individuals. Recent estimates of *Alu* insertion numbers in the human lineage (~7000–9000) suggest that *Alu* elements are amplifying at a rate of one new insert approximately every 15–20 births (see [30] for theory). Thus, it is not surprising that recent *Alu* retrotransposition events have given rise to a number of human diseases.

Alu elements are known to create genetic instability and disease in a number of different ways. We will deal with each mechanism in turn and assess the prevalence, importance and resultant impact on the integrity of the human genome.

Alu Retrotransposition

From a review of current literature, 25 newly integrated *Alu* elements have been determined to induce disease states in human beings. Approximately eleven of the *Alu* elements integrated within introns and either caused partial intron retention within the mature mRNA through *Alu* exonization, or exon skipping [6, 31–36]. A study by Lev-Maor et al. described the process of *Alu* exonization in a 2003 study, where the retention of anti-sense *Alu* elements within the mature mRNA transcript was attributed to the introduction of new splice sites from the *Alu* sequence [32]. One recent study has proposed that exonized *Alu* elements are almost exclusively alternatively spliced, and that ‘*Alu*alternative’ splicing is accountable for producing variable exonic transcripts in over 5% of genes [37]. The retention of *Alu* elements within mRNA transcripts could contribute to subtle differences in gene expression between individuals and populations.

Alu repeats are rarely found within the coding regions of genes, as this may disrupt the gene’s function. However, twelve exon insertion events have been described in the literature (see review [6]). Since the publication of that review in 2001, two other studies have reported *Alu* integration into exons as the cause of genetic disease. In the first case, a young *Alu*Ya5 element inserted into codon 650 of the renal chloride channel gene, *CLCN5*, resulting in Dent’s disease, a cause of renal failure [38]. The second study reports a case of hemophilia A as a direct result of *Alu* integration into exon 14 of the Factor VIII gene [39]. The total number of *Alu* retrotransposition insertions (both intronic and exonic) contributing to disease phenotypes within the human lineage equals 25. The total number of mutations in the Human Gene Mutation Database (<http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html>) currently exceeds 44,000, as of January 2005. Therefore, *Alu* element insertional disruption accounts for 0.05% of all human mutations. However, only non-lethal mutations that cause observable phenotypes will be captured by this statistic. *Alu* insertions that are lethal and those that cause only mild phenotypes will be missed and thereby underestimate the true number of detrimental *Alu* insertions.

Alu-Alu Recombination

Alu-Alu unequal homologous recombination usually involves crossover between evolutionarily older elements within the genome (see [40]). *Alu* elements appear to possess particular characteristics that make them prone to recombination. These are: (1) the relatively close proximity of *Alu* elements within the genome, making most recombination events tolerable. (2) The sequence identity of *Alu* elements (greater than 75%, on average), which promotes efficient base pairing during crossover. (3) The vast number of *Alu* elements that create numerous identical DNA stretches, increasing the probability for recombination. (4) A chi-like motif within the *Alu* sequence that may stimulate recombination. Since 1999, approximately 25 new *Alu-Alu* recombination events have been linked to human disease. This makes the updated contribution of *Alu-Alu* recombination (both germline and somatic) to human genetic disease 0.17% (74/44,000).

Alu elements have also been linked to the presence of gene-rich segmental duplications within the human genome [41]. Given that 5–6% of the human genome sequence was created through segmental duplication events, *Alu-Alu* recombination may have contributed significantly to altered gene expression and species evolution [41]. In addition, mobile element recombination may occur in regions devoid of genes and still impact gene expression [42]. The fact that gene expression can be altered by the recombination of non-coding DNA is especially interesting since it is estimated that over 40 polymorphic *Alu-Alu* recombination events exist within humans (unpublished data). *Alu-Alu* recombination may therefore play a significant role in determining individual- and population-specific disease susceptibility.

Novel Mechanisms of *Alu*-Mediated Genomic Instability

Two novel mechanisms of *Alu*-associated genomic instability have recently been reported, *Alu* retrotransposition-mediated deletion [22] and gene conversion-mediated deletion [43]. Both mechanisms involve the retrotransposition of a new *Alu* element coupled to the deletion of genomic material at the target integration site. *Alu* retrotransposition-mediated deletion involves the integration of an *Alu* cDNA transcript at a new site in the genome, similar to the retrotransposition-mediated deletion mechanism of L1. Gene conversion-mediated deletion involves the non-reciprocal conversion of an older *Alu* element into a younger *Alu* element. Due to the retrotransposition activity of *Alu* elements within humans over the last five million years, numerous chances have arisen for both types of deletion-inducing events.

A recent study of retrotransposition-mediated deletion determined that approximately 9,000 bases of human DNA have been deleted through this

process [22]. In one instance, a 1002-bp deletion caused the functional loss of a retroviral transforming gene, *c-rel*, within the human lineage [22]. Research indicates that *c-rel* may have important roles in regulating cell proliferation and differentiation [44]. If the entire primate order is taken into account, approximately one megabase of DNA may have been deleted through *Alu* retrotransposition-mediated deletion since *Alu* elements evolved 65 million years ago.

Gene conversion-mediated deletion events have yet to be studied in such detail, although preliminary data suggest this mechanism could be as prevalent, if not more, than retrotransposition-mediated deletion (unpublished). The first published example of exonic disruption mediated by gene-conversion deletion occurred in the *CMAH* gene in humans [45]. The deletion event encompassed a 92-bp exon encoding CMP-*N*-acetylneuraminic acid hydroxylase. The partial deletion of *CMAH* induced a biochemical difference in a sialic acid cell surface receptor between humans and non-human primates. Only two other examples of gene conversion-mediated deletion have been reported to date, and arise from the young *Alu*Yg6 and Yb8 subfamilies [43, 46]. Given the fact that *Alu* elements tend to reside in gene rich regions, gene conversion-mediated deletion by young *Alu* family members may be responsible for the deletion of other exonic or regulatory regions within the human genome.

SVA Elements

The SVA element is the least well-documented retrotransposon residing within the human genome. First reported in 1994, SVA elements are a composite retrotransposon consisting of a SINE-R element, a variable number of tandem repeats (VNTR) section and an *Alu* component, all contained within direct repeats (fig. 1c) (see [8]). A recent computational study of SVA elements indicated that there are approximately 1,750–3,500 SVA elements in the human haploid genome, substantially fewer than other retrotransposons such as *Alu* and L1. Low nucleotide sequence divergences within the SVA family suggest that their small number may be the result of their recent proliferation and origin, rather than low retrotranspositional activity. SVA retrotransposition has been verified from studies documenting their involvement in the induction of disease states. Previous research has revealed the presence of an SVA-mediated transduction within the α -spectrin gene (*SPTA1*) [8]. Two other cases of disease-causing SVA insertions have also been reported. The first describes an SVA insertion into an intron of the *BTK* gene, resulting in immunodeficiency X-linked agammaglobulinemia (XLA) [8]. The second case was reported as a cause for Fukuyama-type congenital muscular dystrophy, following disruption of the fukutin gene (see review [8]).

Collectively, L1, *Alu* and SVA retrotransposable elements are responsible for 0.27% (118/44,000) of all human mutations discovered to date. They introduce genetic variation, and disease, on occasion, to human beings via an array of interesting mechanisms. Although researchers in the field of human genetics have explored the major mutational mechanisms of retrotransposable elements, their overall contribution to genomic diversity remains to be quantified.

Acknowledgements

Research on mobile elements in the Batzer laboratory is supported by the Louisiana Board of Regents Millennium Trust Health Excellence Fund HEF (2000-05)-01 (M.A.B.), National Science Foundation BCS-0218338 (M.A.B.) and EPS-0346411 (M.A.B.) and the State of Louisiana Board of Regents Support Fund (M.A.B.).

References

- 1 International Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome. *Nature* 2001;409:860–921.
- 2 Wessler SR: Plant transposable elements. A hard act to follow. *Plant Physiol* 2001;125:149–151.
- 3 Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskaas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, Ashburner M, Celniker SE: The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol* 2002;3:RESEARCH0084.
- 4 Kleckner N: Transposable elements in prokaryotes. *Annu Rev Genet* 1981;15:341–404.
- 5 Batzer MA, Deininger PL: Alu repeats and human genomic diversity. *Nat Rev Genet* 2002;3:370–379.
- 6 Ostertag EM, Kazazian HH Jr: Biology of mammalian L1 retrotransposons. *Annu Rev Genet* 2001;35:501–538.
- 7 Boeke JD: LINEs and Alus – the polyA connection. *Nat Genet* 1997;16:6–7.
- 8 Ostertag EM, Goodier JL, Zhang Y, Kazazian HH Jr: SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am J Hum Genet* 2003;73:1444–1451.
- 9 Orgel LE, Crick FH: Selfish DNA: the ultimate parasite. *Nature* 1980;284:604–607.
- 10 Jelinek WR, Toomey TP, Leinwand L, Duncan CH, Biro PA, Choudary PV, Weissman SM, Rubin CM, Houck CM, Deininger PL, Schmid CW: Ubiquitous, interspersed repeated sequences in mammalian genomes. *Proc Natl Acad Sci USA* 1980;77:1398–1402.
- 11 Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, Taccioli GE, Batzer MA, Moran JV: DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat Genet* 2002;31:159–165.
- 12 Chu WM, Ballard R, Carpick BW, Williams BR, Schmid CW: Potential Alu function: regulation of the activity of double-stranded RNA-activated kinase PKR. *Mol Cell Biol* 1998;18:58–68.
- 13 Myers JS, Vincent BJ, Udall H, Watkins WS, Morrish TA, Kilroy GE, Swergold GD, Henke J, Henke L, Moran JV, Jorde LB, Batzer MA: A comprehensive analysis of recently integrated human Ta L1 elements. *Am J Hum Genet* 2002;71:312–326.
- 14 Martinez-Garay I, Ballesta MJ, Oltra S, Orellana C, Palomeque A, Molto MD, Prieto F, Martinez F: Intronic L1 insertion and F268S, novel mutations in *RPS6KA3* (RSK2) causing Coffin-Lowry syndrome. *Clin Genet* 2003;64:491–496.
- 15 van den Hurk JA, van de Pol DJ, Wissinger B, van Driel MA, Hoefsloot LH, de Wijs IJ, van den Born LI, Heckenlively JR, Brunner HG, Zrenner E, Ropers HH, Cremers FP: Novel types of mutation in the choroideremia (*CHM*) gene: a full-length L1 insertion and an intronic mutation activating a cryptic exon. *Hum Genet* 2003;113:268–275.

- 16 Mukherjee S, Mukhopadhyay A, Banerjee D, Chandak GR, Ray K: Molecular pathology of haemophilia B: identification of five novel mutations including a LINE 1 insertion in Indian patients. *Haemophilia* 2004;10:259–263.
- 17 Han JS, Szak ST, Boeke JD: Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* 2004;429:268–274.
- 18 Perepelitsa-Belancio V, Deininger P: RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nat Genet* 2003;35:363–366.
- 19 Kazazian HH Jr, Goodier JL: LINE drive. Retrotransposition and genome instability. *Cell* 2002;110:277–280.
- 20 Gilbert N, Lutz-Prigge S, Moran JV: Genomic deletions created upon LINE-1 retrotransposition. *Cell* 2002;110:315–325.
- 21 Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD: Human L1 retrotransposition is associated with genetic instability in vivo. *Cell* 2002;110:327–338.
- 22 Callinan PA, Jianxin W, Herke SW, Garber RK, Liang P, Batzer MA: *Alu* retrotransposition-mediated deletion. *J Mol Biol* 2005;348:791–800.
- 23 Narita N, Nishio H, Kitoh Y, Ishikawa Y, Minami R, Nakamura H, Matsuo M: Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. *J Clin Invest* 1993;91:1862–1867.
- 24 Kondo-Iida E, Kobayashi K, Watanabe M, Sasaki J, Kumagai T, Koide H, Saito K, Osawa M, Nakamura Y, Toda T: Novel mutations and genotype-phenotype relationships in 107 families with Fukuyama-type congenital muscular dystrophy (*FCMD*). *Hum Mol Genet* 1999;8:2303–2309.
- 25 Ho HJ, Ray DA, Salem AH, Myers JS, Batzer MA: Straightening out the LINES: LINE-1 orthologous loci. *Genomics* 2005;85:201–207.
- 26 Vincent BJ, Myers JS, Ho HJ, Kilroy GE, Walker JA, Watkins WS, Jorde LB, Batzer MA: Following the LINES: an analysis of primate genomic variation at human-specific LINE-1 insertion sites. *Mol Biol Evol* 2003;20:1338–1348.
- 27 Holmes SE, Dombroski BA, Krebs CM, Boehm CD, Kazazian HH Jr: A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. *Nat Genet* 1994;7:143–148.
- 28 Moran JV, DeBerardinis RJ, Kazazian HH Jr: Exon shuffling by L1 retrotransposition. *Science* 1999;283:1530–1534.
- 29 Goodier JL, Ostertag EM, Kazazian HH Jr: Transduction of 3'-flanking sequences is common in L1 retrotransposition. *Hum Mol Genet* 2000;9:653–657.
- 30 Deininger PL, Batzer MA: Evolution of retrotransposons; in Hecht MK (ed.): *Evolutionary Biology*. Plenum Press, New York, 1993, pp 157–196.
- 31 Mitchell GA, Labuda D, Fontaine G, Saudubray JM, Bonnefont JP, Lyonnet S, Brody LC, Steel G, Obie C, Valle D: Splice-mediated insertion of an *Alu* sequence inactivates ornithine delta-amino-transferase: a role for *Alu* elements in human mutation. *Proc Natl Acad Sci USA* 1991;88:815–819.
- 32 Lev-Maor G, Sorek R, Shomron N, Ast G: The birth of an alternatively spliced exon: 3' splice-site selection in *Alu* exons. *Science* 2003;300:1288–1291.
- 33 Ganguly A, Dunbar T, Chen P, Godmilow L, Ganguly T: Exon skipping caused by an intronic insertion of a young *Alu* Yb9 element leads to severe hemophilia A. *Hum Genet* 2003;113:348–352.
- 34 Knebelmann B, Forestier L, Drouot L, Quinones S, Chuet C, Benessy F, Saus J, Antignac C: Splice-mediated insertion of an *Alu* sequence in the *COL4A3* mRNA causing autosomal recessive Alport syndrome. *Hum Mol Genet* 1995;4:675–679.
- 35 Ferlini A, Galie N, Merlini L, Sewry C, Branzi A, Muntoni F: A novel *Alu*-like element rearranged in the dystrophin gene causes a splicing mutation in a family with X-linked dilated cardiomyopathy. *Am J Hum Genet* 1998;63:436–446.
- 36 Vervoort R, Gitzelmann R, Lissens W, Liebaers I: A mutation (IVS8+0.6kbdelTC) creating a new donor splice site activates a cryptic exon in an *Alu*-element in intron 8 of the human beta-glucuronidase gene. *Hum Genet* 1998;103:686–693.
- 37 Kreahling J, Graveley BR: The origins and implications of *Alu* alternative splicing. *Trends Genet* 2004;20:1–4.

- 38 Claverie-Martin F, Gonzalez-Acosta H, Flores C, Anton-Gamero M, Garcia-Nieto V: De novo insertion of an *Alu* sequence in the coding region of the *CLCN5* gene results in Dent's disease. *Hum Genet* 2003;113:480–485.
- 39 Sukarova E, Dimovski AJ, Tchacarova P, Petkov GH, Efremov GD: An *Alu* insert as the cause of a severe form of hemophilia A. *Acta Haematol* 2001;106:126–129.
- 40 Deininger PL, Batzer MA: *Alu* repeats and human disease. *Mol Genet Metab* 1999;67:183–193.
- 41 Bailey JA, Liu G, Eichler EE: An *Alu* transposition model for the origin and expansion of human segmental duplications. *Am J Hum Genet* 2003;73:823–834.
- 42 Balemans W, Patel N, Ebeling M, Van Hul E, Wuyts W, Lacza C, Dioszegi M, Dikkers FG, Hildering P, Willems PJ, Verheij JB, Lindpaintner K, Vickery B, Foerzler D, Van Hul W: Identification of a 52 kb deletion downstream of the *SOST* gene in patients with van Buchem disease. *J Med Genet* 2002;39:91–97.
- 43 Salem AH, Kilroy GE, Watkins WS, Jorde LB, Batzer MA: Recently integrated *Alu* elements and human genomic diversity. *Mol Biol Evol* 2003;20:1349–1361.
- 44 Bishop JM, Varmus H: Functions and origins of retroviral oncogenes; in Weiss R, Teich N, Varmus H, Coffin J (eds): *Molecular Biology of the Tumor Viruses: RNA Tumor Viruses*. Cold Spring Harbor, New York, 1982, pp 999–1108.
- 45 Hayakawa T, Satta Y, Gagneux P, Varki A, Takahata N: *Alu*-mediation inactivation of the human CMP-N-acetyl-neuraminic acid hydroxylase gene. *Proc Natl Acad Sci USA* 2001;98:11399–11404.
- 46 Carter AB, Salem A-H, Hedges DJ, Nguyen Keegan C, Kimball B, Walker JA, Watkins WS, Jorde LB, Batzer MA: Genome wide analysis of the human Yb lineage. *Hum Genomics* 2004;1:167–168.

Mark A. Batzer

Department of Biological Sciences

Biological Computation and Visualization Center

Center for BioModular Multi-Scale Systems

Louisiana State University, 202 Life Sciences Building

Baton Rouge, LA 70803 (USA)

Tel. +1 225 578 7102, Fax +1 225 578 7113, E-Mail mbatzer@lsu.edu