Molecular Genetics and Metabolism **67,** 183–193 (1999)
Article ID mgme.1999.2864, available online at http://www.idealibrary.com on IDEAL®

MINIREVIEW

Alu Repeats and Human Disease

Prescott L. Deininger*,†,[1] and Mark A. Batzer‡

*Tulane Cancer Center, SL-66, and Department of Environmental Health Sciences, Tulane University Medical Center, 1430 Tulane Avenue, New Orleans, Louisiana 70112; †Laboratory of Molecular Genetics, Ochsner Medical Foundation, 1516 Jefferson Highway, New Orleans, Louisiana 70121; and ‡Departments of Pathology , Biochemistry and Molecular Biology Biometry and Genetics, Stanley S. Scott Cancer Center, Neuroscience Center of Excellence, Louisiana State University Medical Center, 1901 Perdido Street, New Orleans, Louisiana 70112

Received March 25, 1999, and in revised form March 31, 1999

Alu elements have amplified in primate genomes through a RNA-dependent mechanism, termed retroposition, and have reached a copy number in excess of 500,000 copies per human genome. These elements have been proposed to have a number of functions in the human genome, and have certainly had a major impact on genomic architecture. Alu elements continue to amplify at a rate of about one insertion every 200 new births. We have found 16 examples of diseases caused by the insertion of Alu elements, suggesting that they may contribute to about 0.1% of human genetic disorders by this mechanism. The large number of Alu elements within primate genomes also provides abundant opportunities for unequal homologous recombination events. These events often occur intrachromosomally, resulting in deletion or duplication of exons in a gene, but they also can occur interchromosomally, causing more complex chromosomal abnormalities. We have found 33 cases of germline genetic diseases and 16 cases of cancer caused by unequal homologous recombination between Alu repeats. We estimate that this mode of mutagenesis accounts for another 0.3% of human genetic diseases. Between these different mechanisms, Alu elements have not only contributed a great deal to the evolution of the genome but also continue to contribute to a significant portion of human genetic diseases.

[1] To whom correspondence should be addressed at Tulane Cancer Center, SL-66, Tulane University Medical Center, 1430 Tulane Ave., New Orleans, LA 70112. Fax: (504) 588-5516. E-mail: pdeinin@tcs.tulane.edu.

Key Words: **Alu repeats; recombination; insertion mutation; human disease; genetic diversity.**

THE SPREAD OF Alu ELEMENTS IN THE HUMAN GENOME

Alu elements represent a sequence of approximately 300 nucleotides (nt) in length that are transcribed by RNA polymerase III. The RNA transcript is then reverse-transcribed and inserted into a new location in the genome. This RNA-mediated process for making new copies of the element is termed retroposition (1). Different Alu elements in the genome are not identical to one another. It appears that Alu elements that have integrated recently within the genome are quite homogeneous, and almost exact copies of one another (2). However, the older copies have accumulated random mutations, making them typically divergent by 20% or more from one another at the sequence level (3).

Alu elements began inserting early in primate evolution, approximately 65 mya (3). Although there are some related elements in mammals outside of the primate order, they do not have the specific structure of Alu elements. The rate of Alu amplification appears to have reached a maximum between 35 and 60 mya, and is currently amplifying at only 1% of the maximum rate. There are probably only about 2000 Alus specific to the human genome, and not found in chimpanzee and gorilla. Thus, about 99.8% of the 500,000 Alus in the human genome can

1096-7192/99 $30.00

TABLE 1
Alu Insertions and Disease

Locus	Distribution	Subfamily	Disease	Reference
CaR	Familial	Ya4	Hypocalciuric hypercalcemia and neonatal severe hyperparathyroidism	(51)
Mlvi-2	*De novo* (somatic?)	Ya5	Associated with leukemia	(52)
NF1	*De novo*	Ya5	Neurofibromatosis	(53)
PROGINS	About 50%	Ya5	Linked with ovarian carcinoma	(54)
IL2RG	Familial	Ya5	XSCID	(55)
ACE	About 50%	Ya5	Linked with protection from heart disease	(35)
Factor IX	A grandparent	Ya5	Hemophilia	(56)
EYA1	*De novo*	Ya5	Branchio-oto-renal syndrome	(57)
2 × FGFR2	*De novo*	Ya5 & Yb8	Apert's syndrome	(41)
Cholinesterase	One Japanese family	Yb8	Cholinesterase deficiency	(58)
APC	Familial	Yb8	Hereditary desmoid disease	(59)
Btk	Familial	Y	X-linked agammaglobulinaemia	(55)
C1 inhibitor	*De novo*	Y	Complement deficiency	(60)
BRCA2	*De novo*	Y	Breast cancer	(61)
GK	?	Y	Glycerol kinase deficiency	(62)

be found at the same locus in all of the great apes, and 85% of the elements at specific loci can be found in all monkeys. Our best estimates of Alu amplification in the human genome are that there is one new insert in about every 200 new births (4). Although this is well below the peak rate, it is still high enough to represent a significant factor in human mutagenesis.

In addition to random mutations, which occur to Alu elements after their insertion in the genome, there are specific base changes that allow separation of Alu elements into different subfamilies (5–10). The different subfamilies were all inserted at different stages of primate evolution. Almost all of the insertions that have occurred specifically in the human genome come from four closely related subfamilies, Alu Y, Ya5, Ya8, and Yb8. Ya5 and Yb8 inserts represent the majority of the inserts and Alu Y inserts are relatively rare. All of the new inserts belong to a small group of the most recently created subfamilies (see Table 1). This demonstrates that only a small subset of Alus is capable of amplification (11).

Several explanations for the selective amplification of specific subfamilies have been proposed. One likely explanation is that a few specific loci are capable of active amplification, while almost all other loci are not, and that there are almost no such loci in the older subfamilies (11). Alternatively, one has to propose that loci from all subfamilies express, but that the RNAs expressed from the newer subfamilies interact with the retroposition apparatus much better than the older subfamily RNAs (12,13).

Alus AND L1 ELEMENTS

The other major mobile element in the human genome is the L1 element. Alu elements are RNA polymerase III-derived transcripts that have no coding capacity. Thus, they do not code for any proteins that might be involved in the retroposition process. L1 repeats, on the other hand, are much longer and have two open-reading frames (reviewed in (14)). One open-reading frame apparently codes for an RNA-binding protein whose exact function is unknown. The other open-reading frame codes for a protein that includes domains for reverse transcriptase, as well as for an endonuclease that apparently nicks the genome at the site of insertion (15–17). An assay that allows rapid L1 retroposition in cultured cells has been devised recently (18). This assay facilitates the dissection of the details of the L1 retroposition mechanism.

Alu elements must obtain the enzymes for their retroposition from somewhere. In addition, there are striking similarities between the mechanisms of Alu and L1 retroposition that make it very attractive to think that L1 elements may supply the necessary components for Alu retroposition (15,16,19,20). This idea is certainly very attractive, and thus the rate of Alu retroposition may be very dependent on the rate and evolution of L1 elements.

Alu ELEMENTS: FUNCTIONAL ROLE OR A PARASITE'S PARASITE

Alu repeats represent over 5% of the mass of the human genome. They are also spread throughout the entire genome, at varying densities. These observations, along with other specific properties of the Alu elements have led to a number of hypothetical functions for the Alu elements that might explain their ubiquitous presence in primate genomes. Some of the proposed roles involve an everyday function for the cell, while others are of a more sporadic nature.

The first role ever proposed for Alu elements was that they might be origins of DNA replication (21). This role is consistent with their high copy number and dispersed nature, but has not been substantiated by direct experimentation and seems like too important a function to be served by an element that is not found outside of primates.

More recently, evidence has been presented that Alu RNAs may stimulate protein translation by inhibiting a RNA-dependent protein kinase, PKR (22–24). Because Alu RNAs from many loci are stimulated by a number of cellular stresses, such as viral infection and heat shock, this would provide a mechanism by which dispersed sequences may contribute to a cellular process as a group. If this is a function of Alu elements, then it is likely to represent only a slightly modified regulation seen in nonprimate species that is filled by other RNAs or molecules in those species.

Evidence has been presented in yeast that retrotransposable elements may aid in healing chromosomal breaks (25,26). This suggests the possibility that Alu and L1 elements may provide the same role in the human genome.

There are several thoughts concerning the possible roles of Alu elements in the evolution of the human genome. As discussed below, Alu elements can lead to unequal recombination that results in deletion or duplication of sequences. These events could allow duplication of exons and therefore formation of new protein variants. They can also contribute to interchromosomal recombination that may lead to cytogenetic alterations that are involved in human speciation.

There are also several ways in which Alu repeats have been proposed to influence the evolution of gene expression. Because Alu elements are rich in CpG dinucleotides that represent the substrate for genomic methylation, Alu elements represent CpG-rich islands that make up about 30% of the methylation sites in the human genome (24). When an Alu element inserts in a new location in the genome, it introduces a CpG island at that new location. CpG islands have been associated with gene regulation, as well as imprinting of genes, and therefore Alu elements may contribute to the evolution of gene expression and imprinting in the human genome. In addition, Alu elements have been found to carry functional promoter elements for several of the steroid hormone receptors (27,28). Thus, insertion of a new Alu element in the vicinity of a gene may introduce new transcription factor-binding sites that could alter the regulation of gene expression. There are a number of cases where elements that influence gene expression have been mapped to within an Alu repeat (29), demonstrating that the introduction of these sequences can at least occasionally contribute to gene expression and regulation.

Although, there are numerous cases where individual Alu elements have had a positive impact on the human genome, it might be argued that none of them has been confirmed as a function. In this sense we would not define something that happens in a positive sense every few thousand years as being a function, because it would be occurring too sporadically to apply a positive selection for the presence of Alu elements. In addition, studies of individual Alu elements demonstrate that there is essentially no selective pressure on any given Alu repeat, although it is possible that selection does exist for a handful of master elements. Thus, it has been argued that Alu and L1 elements may both represent "selfish" DNA, or DNA that is only working to replicate itself. Selfish DNA may often have negative impacts on the host, but can be tolerated if it does not have too strong an adverse affect. Selfish DNA may also occasionally have positive benefits, but only by chance, and not by functional design. If L1 elements are essentially a parasite within the human genome, and if Alu relies on L1 elements for their amplification process, then one might describe Alu as a "parasite's parasite."

Alus AS MARKERS FOR HUMAN DIVERSITY

Although there is still a question as to whether there is a true functional role for Alu elements in the human genome, Alu elements have proved to be

useful in studies of human DNA. The presence of Alu repeats located ubiquitously throughout the human genome, but not in nonprimate species, has allowed detection of human DNA sequences that have been transfected into the cells of other organisms, such as mice. This has been useful in marker-rescue experiments in isolating a number of genes, including the first examples of oncogenes isolated by transforming rodent cell lines with human tumor DNAs (30). More recently, inter-Alu PCR (31,32) has found a broad range of uses in isolating specific human DNA regions from mouse/human hybrid cell lines and other complex sources containing large segments of human DNA.

Recent Alu insertions have also proven useful in a number of human population studies. In particular, there are over 1000 Alu insertions that occurred recently enough to be present only in a subset of human chromosomes. Because there does not seem to be any specific mechanism for removing Alu elements from the genome, once inserted they make a very stable genetic marker (33,34). This observation, along with the extremely low probability that any two recently integrated elements have inserted independently in the same chromosomal location, makes Alu insertions one of the best identical-by-descent (IBD) markers for human evolution studies. Any two individuals sharing an Alu insert almost certainly do so because they share a common ancestor in which the insertion occurred. Table 1 includes an example of an Alu insertion in the angiotensin-converting enzyme (ACE) locus that shows a useful association with protective advantages from heart disease (35). Many other Alu insertion polymorphisms have been identified either in random genomic loci or in specific genes, but without any known disease association. These Alu insertions are easy to assay for their presence or absence in a chromosomal location and have been found to be very powerful markers for human forensic and molecular anthropology studies (36,37).

RETROPOSITION OF Alu ELEMENTS AND DISEASE

Alu elements are located throughout the genome and in almost any location within a gene except those in which they would totally disrupt the function of that gene. Figure 1 illustrates some of the positions relative to a typical gene structure in which Alu may land. Alus landing far enough upstream of a gene may have no influence on that gene's expression. However, Alus landing in or near the promoter/enhancer regions of a gene have been found to influence the expression of specific genes (reviewed in (29)), as well as to have the general potential to add transcription elements, like steroid hormone receptor elements (27,28), to the upstream gene region.

Very few Alu elements are found within the 5′ noncoding or coding regions of exons, presumably because insertions in those locations are too disruptive to gene function. There are a number of instances where Alu elements have been found to be part of the region coding for the carboxy-terminus of a protein product (38,39). Presumably these Alus insert far enough downstream in the coding sequence to result in a new carboxy-terminus that does not disrupt the structure of the protein.

Insertions into the 3′ noncoding regions of genes are found commonly and appear to have few negative affects. Similarly Alus are commonly found in introns, demonstrating that Alu insertions in much of the intronic region do not alter gene function significantly.

The vast majority of Alu insertions that have led to human disease insert into coding exons, or into introns relatively near an exon and presumably alter splicing. Table 1 is a list of the genetic defects that are thought to be caused by Alu insertion events. Not all of these cases have been demonstrated to be directly causative for the disease, but the rarity of Alu insertion events, coupled with the lack of other detectable mutations in these cases, strongly indicates that these are the causative events. The ACE insertion (35,40) is likely to be one example, however, that shows association with disease, but is highly unlikely to be the causative event.

The above examples demonstrate that Alu insertions are capable of causing genetic defects which lead to human disease. Examples of this type are being found at an increasing frequency as the tools for genetic analysis allow more mutations to be detected. Finding 16 Alu-based insertion mutations in the Human Genetic Mutation Database that contains 14374 characterized human mutations suggests that Alu elements contribute to approximately 0.1% of human genetic diseases. This number agrees well with a previous calculation based on a similar dataset of mutations where Alu and L1 insertions were estimated to each contribute approximately 0.075% of human mutations (16). In some cases, the insertional mutagenesis may make detection of mutations easier, biasing the results in favor of the

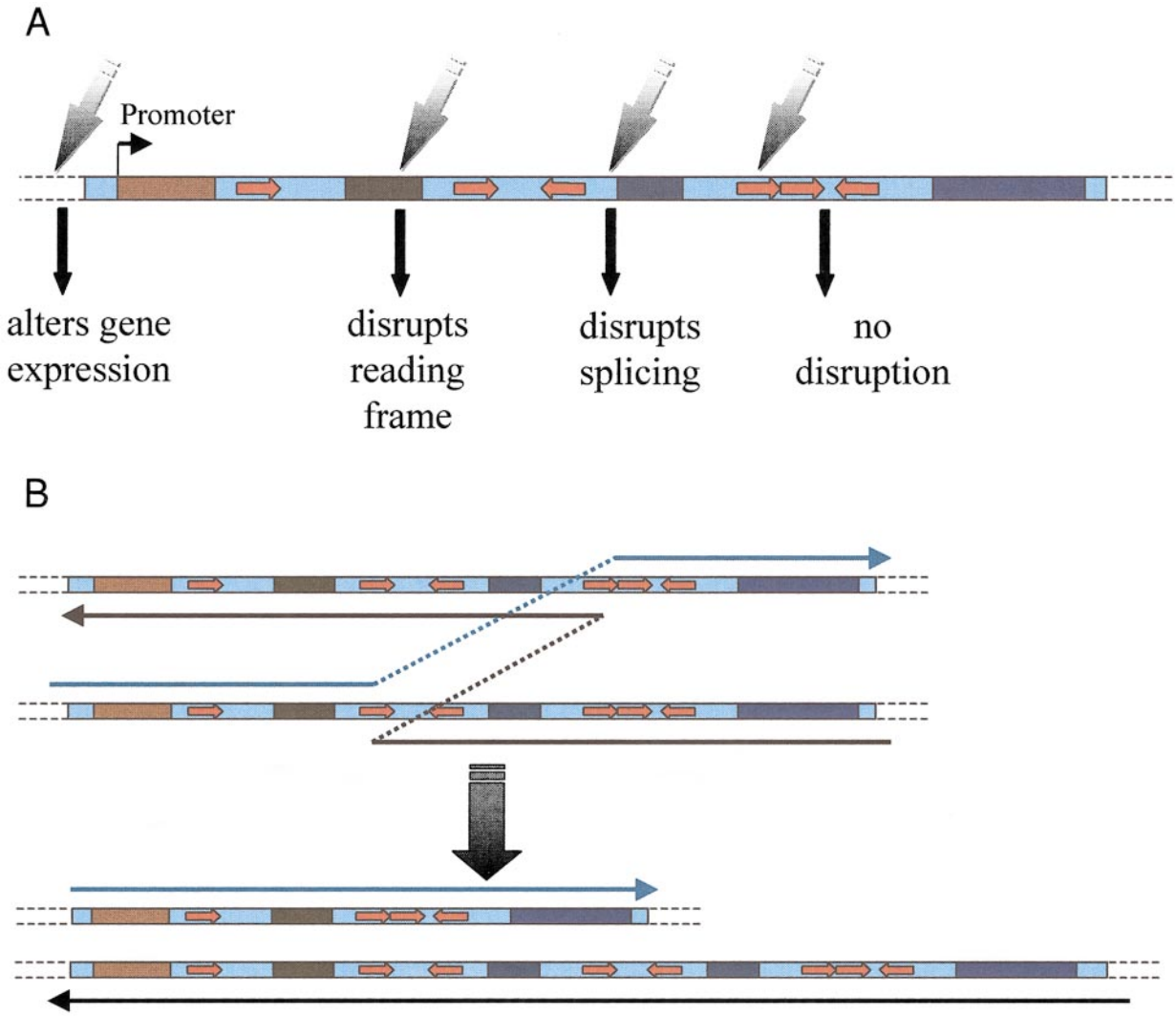

FIG. 1. Schematic of Alu-induced damage to the human genome. Panel A illustrates some of the potential consequences of insertion of a new element in the vicinity of a gene. The colored boxes represent various exons of the gene. The red arrows show existing Alu elements oriented in different directions in the introns of the gene. Depending on the site of insertion, the Alu element has varied probability of impact on the genome as shown. Panel B illustrates an unequal, homologous recombination occurring between two Alu elements in different introns of a gene. The arrows broken by dotted lines show the path of the recombination event. The genes below show that one copy will have a deletion while the other will duplicate gene sequences. Either is likely to be deleterious.

detection of Alu insertions. However, many mutation detection strategies are designed to identify point mutations, particularly in coding regions, and may overlook insertions, particularly if they occur in introns. In addition, many new mobile element insertions may be lethal during embryogenesis. Therefore, it is likely that these estimates of insertion frequencies are underestimates of the true contribution of new Alu insertions to human disease.

We expect that with increasing study of mutations, it will be found that some genetic diseases are more likely than others to result from retroposon insertion. It has certainly been observed that some genes have a much higher Alu repeat content, making it reasonable that they will have a higher frequency of disabling Alu insertions. It has been observed that 2 out of 258 mutations in the FGFR2 gene were caused by Alu insertions (41). This is the first case of multiple Alu insertion mutations being detected associated with a single disease, suggesting that this genetic locus may be more susceptible to retroposon insertions than other regions of the ge-

TABLE 2
Alu/Alu Recombination and Germ-Line Disease

Locus	Distribution	Disease	Reference
8 × LDLR	Kindreds	Hypercholesterolemia	(63–67)
5 × α-globin	Kindreds	α-thalassaemia	(68–71)
5 × C1 inhibitor	Kindred	Angioneurotic adema	(60,72)
Lys Hydrox.	Kindreds	Ehlers-Danlos syndrome	(73)
DMD	Kindred	Duchenne's muscular dystropy	(74)
ADA	One patient	ADA deficiency-SCID	(75)
Apo B	One patient	Hypo-betalipoproteinemia	(76)
Ins. Rec. β	One patient	Insulin-independent diabetes	(77)
α-gal A	One patient	Fabry disease	(78)
HPRT	One patient	Lesch-Nyhan syndrome	(79)
Plat. Fibrinogen Receptor	Kindred	Glanzmann thrombasthenia	(80)
Phosphorylase kinase	One patient	Glycogen storage disease	(81)
GALNS	One patient	Mucopolysaccharidosis type IVA	(82)
Antithrombin	One patient	Thrombophilia	(83)
XY	One patient	XX male	(84)
β-HEXA	Classic form of disease	Tay Sachs	(85)
C3	Kindred	C3 deficiency	(86)
HEXB	27% of patients	Sandhoff's disease	(87)

nome. However, the number of insertions found so far is still fairly low making more definitive conclusions difficult.

RECOMBINATION BETWEEN Alu ELEMENTS ASSOCIATED WITH DISEASE

In addition to the potential impact of Alu element insertions in causing human disease, their dispersion throughout the genome provides ample opportunity for unequal homologous recombination which leads to a much higher level of mutations. Figure 1B illustrates how this unequal recombination can cause insertion or deletion mutations. When recombination occurs between Alu elements on the same chromosome, the result is that there is either duplication or deletion of the sequences between the Alus. Recombination may also occur between Alu elements on different chromosomes, resulting in chromosomal translocations or more complex chromosomal rearrangements.

Table 2 presents a compilation of Alu/Alu recombination events that have contributed to germ-line disease with Alu-based recombination events associated with cancer shown in Table 3. There are many more recombination than insertion events contributing to disease and the table of recombination events is not intended to be exhaustive in presenting all of the Alu/Alu recombinations that have contributed to human disease. In addition, there are many recombination events that occurred between an Alu element and some other non-Alu-related sequence which may have been influenced by the presence of the Alu element (42). Although single Alu elements may contribute specifically to such recombination events, we have made no efforts to collect those data. The mutations resulting from Alu/Alu recombination include 33 mutations that are the result of germ-line recombination and 16 mutations that are the result of somatic events that led to cancer. Based on the calculations in the previous section, the germline recombination mutants would represent about 0.3% of mutants characterized. We expect that this number is an underestimate as mutation schemes aimed at detecting point mutants would often be expected to overlook large duplication and deletion events, and we have probably not reported all known Alu/Alu recombinations in the tables.

The data in Tables 2 and 3 show that Alu/Alu recombination events are highly biased towards specific genes. The first to show evidence for this was the LDLR gene, which has at least eight independent cases. It was also reported that these recombination events appeared to take place in a preferred location within the Alu element (42,43). These data suggested that Alu elements may represent hot spots for recombination by a mechanism that was more than simple homologous recombination. Multiple Alu/Alu recombination events have also occurred in the germ line involving two other genes.

TABLE 3
Alu/Alu Recombination and Cancer

Locus	Distribution	Disease	Reference
10 × ALL-1 (MLL)	Somatic	Acute myelogenous leukemia	(88–90)
2 × BRCA1	Somatic and kindreds	Breast cancer	(91,92)
MLH1	Two kindreds	HNPCC	(93)
TRE	Somatic	Ewing's sarcoma	(94)
RB	Common	Association with glioma	(95)
EWS	Subset of Africans	Protective against Ewing sarcoma?	(96)

Even more striking is the preferential recombination seen in somatic recombination. The All-1 gene which participates in a high proportion of acute leukemias is another hotspot for Alu/Alu recombination. This includes intragenic recombination which is the major cause of acute myelogenous leukemia in individuals without a cytogenetic defect, as well as a possible contribution to recombination between the All-1 gene and other chromosomal loci in causing more complex cytogenetic defects associated with leukemia (44–46).

The genes that show high levels of Alu/Alu recombination tend to have a large number of Alu sequences. Although Alu density may help contribute to this recombination, the correlation does not seem to hold up upon analysis of other Alu-rich genes. Therefore, it seems likely that some other factor contributes to the high recombination rates seen in these genes and that the Alu elements are likely to help in that process rather than to be the primary cause.

It has generally been found that longer stretches of sequence identity allow more efficient homologous recombination and that 300 bp of imperfect sequence identity would represent a relatively inefficient target (47). Therefore, as Alu elements accumulate random mutations after integration in the genome their recombination potential gradually decreases. Thus, early in primate evolution when a high proportion of Alu elements were closer matches to one another, Alu/Alu recombination may have contributed even more to the evolution and reshaping of primate genomes.

Based on the above considerations, one might expect the much longer L1 family of elements to contribute significantly to recombination, as well. Surprisingly, we are familiar with only two L1/L1 recombination events in the human genome (48). Therefore, it would appear that: (1) L1 elements are located in less recombinogenic regions of the human genome; (2) the approximately 10-fold lower copy number of L1 elements is more than enough to offset their larger size in terms of probabilities of recombination; (3) some basic property of the Alu elements themselves makes them recombinogenic; or (4) the larger average spacing between L1 elements causes the vast majority of L1/L1 recombination events to be lethal. It is possible that all of these factors may contribute to this observed difference. Transient transfection experiments suggest that the third possibility may not be true since Alu sequences did not recombine more frequently than other control sequences (49). However, in their native chromatin environment, or in specific cell types or cell stimuli *in vivo,* Alus may still respond with higher recombination rates. We believe that the fourth possibility may be the dominant factor, however. The vast majority of Alu/Alu recombination events listed in the tables represent recombination between Alu elements within the same gene. This limits the effect of the recombination to a single gene defect. With their lower copy number and tendency to be located between genes rather than in genes, L1/L1 recombination events are likely either to involve only intergenic regions or to involve a much larger region that may cause defects in several genes simultaneously, resulting in loss of viability.

There is growing evidence that repetitive DNAs contribute to disease either through the mutations they cause during the retroposition process that forms them (16,50) or through recombination processes involving unequal cross-overs of repetitive elements. These recombination events may involve repetitive sequences of various repetition frequencies with the likelihood that longer and more perfect repeats that are near one another probably recombine well, while short, mismatched repeats (like Alu) recombine relatively poorly. However, the extremely high copy number of Alu elements makes them a

major factor in the molecular basis of human diseases.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health RO1 GM45668 to P.L.D., Department of the Army DAMD17-98-1-8119 to P.L.D. and M.A.B., and National Science Foundation SBR-9610147 to M.A.B.

REFERENCES

1. Rogers J. Retroposons defined. *Nature* **301:**460, 1998.
2. Batzer M, Kilroy G, Richard P, Shaikh T, Desselle T, Hoppens C, Deininger P. Structure and variability of recently inserted Alu family members. *Nucleic Acids Res* **18:**6793–6798, 1990.
3. Shen M, Batzer M, Deininger P. Evolution of the master Alu gene(s). *J Mol Evol* **33:**311–320, 1991.
4. Deininger P, Batzer M. SINE master genes and population biology. In The Impact of Short, Interspersed Elements (SINEs) on the Host Genome (Maraia R, Ed.). Georgetown, TX: Landes, pp 43–60, 1995.
5. Batzer M, Deininger P. A human-specific subfamily of *Alu* sequences. *Genomics* **9:**481–487, 1991.
6. Jurka J, Smith T. A fundamental division in the Alu family of repeated sequences. *Proc Natl Acad Sci USA* **85:**4775–4779, 1988.
7. Matera AG, Hellmann U, Schmid CW. A transpositionally and transcriptionally competent Alu subfamily. *Mol Cell Biol* **10:**5424–5432, 1990.
8. Quentin Y. The Alu family developed through successive waves of fixation closely connected with primate lineage history. *J Mol Evol* **27:**194–199, 1988.
9. Slagel V, Flemington E, Traina-Dorge V, Bradshaw H, Deininger P. Clustering and sub-family relationships of the Alu family in the human genome. *Mol Biol Evol* **4:**19–29, 1987.
10. Willard C, Nguyen HT, Schmid CW. Existence of at least three distinct Alu subfamilies. *J Mol Evol* **26:**180–186, 1987.
11. Deininger P, Batzer M, Hutchison C, Edgell M. Master genes in mammalian repetitive DNA amplification. *Trends Genet* **8:**307–312, 1992.
12. Sinnett D, Richer C, Deragon JM, Labuda D. Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences. *J Mol Biol* **226:**689–706, 1992.
13. Kim J, Kass DH, Deininger PL. Transcription and processing of the rodent ID repeat family in germline and somatic cells. *Nucleic Acids Res* **23:**2245–2251, 1995.
14. Boeke JD. LINEs and Alus—The polyA connection. *Nature Genet* **16:**6–7, 1997.
15. Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87:**905–916, 1996.
16. Kazazian HH, Jr, Moran JV. The impact of L1 retrotransposons on the human genome. *Nature Genet* **19:**19–24, 1998.
17. Mathias SL, Scott AF, Kazazian HH Jr, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science* **254:**1808–1810, 1991.
18. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. High frequency retrotransposition in cultured mammalian cells. *Cell* **87:**917–927, 1996.
19. Almenoff JS, Jurka J, Schoolnik GK. Induction of heat-stable enterotoxin receptor activity by a human Alu repeat. *J Biol Chem* **269:**16610–16617, 1994.
20. Jurka J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc Natl Acad Sci USA* **94:**1872–1877, 1997.
21. Jelinek W, Toomey T, Leinwand L, Duncan CH, Biro PA, Choudary PV, Weissman S, Rubin C, Houck C, Deininger PL, Schmid CW. Ubiquitous, interspersed repeated sequences in mammalian genomes. *Proc Natl Acad Sci USA* **77:**1398–1402, 1980.
22. Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM, Schmid CW, Zietkiewicz E, Zuckerkandl E. Standardized nomenclature for *Alu* repeats. *J Mol Evol* **42:**3–6, 1996.
23. 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* **18:**58–68, 1998.
24. Schmid CW. Does SINE evolution preclude alu function? *Nucleic Acids Res* **26:**4541–4550, 1998.
25. Garfinkel DJ. Genetic loose change: How retroelements and reverse transcriptase heal broken chromosomes. *Trends Microbiol* **5:**173–175, 1997.
26. Moore JK, Haber JE. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. *Nature* **383:**644–646, 1996.
27. Vansant G, Reynolds WF. The consensus sequence of a major Alu subfamily contains a functional retinoic acid response element. *Proc Natl Acad Sci USA* **92:**8229–8233, 1995.
28. Norris J, Fan D, Aleman C, Marks JR, Futreal PA, Wiseman RW, Iglehart JD, Deininger PL, McDonnell DP. Identification of a new subclass of Alu DNA repeats which can function as estrogen receptor-dependent transcriptional enhancers. *J Biol Chem* **270:**22777–22782, 1995.
29. Britten RJ. DNA sequence insertion and evolutionary variation in gene regulation. *Proc Natl Acad Sci USA* **93:**9374–9377, 1996.
30. Shih C, Weinberg RA. Isolation of a transforming sequence from a human bladder carcinoma cell line. *Cell* **29:**161–169, 1982.
31. Nelson DL, Ballabio A, Victoria MF, Pieretti M, Bies RD, Gibbs RA, Maley JA, Chinault AC, Webster TD, Caskey CT. Alu-primed polymerase chain reaction for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: Identification of clones associated with a disease locus. *Proc Natl Acad Sci USA* **88:**6157–6161, 1991.
32. Ledbetter SA, Nelson DL, Warren ST, Ledbetter DH. Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. *Genomics* **6:**475–481, 1990.

33. Batzer M, Arcot S, Phinney J, Alegria-Hartman M, Kass D, Milligan S, Kimpton C, Gill P, Hochmeister M, Ioannou P, Herrera R, Boudreau D, Scheer WD, Keats B, Deininger P, Stoneking M. Genetic variation of recent *Alu* insertions in human populations. *J Mol Evol* **42:**22–29, 1996.

34. Perna N, Batzer M, Deininger P, Stoneking M. Alu insertion polymorphism: A new type of marker for human population studies. *Hum Biol* **64:**641–648, 1992.

35. Cambien F, Poirier O, Lecerf L, Evans A, Cambou JP, Arveiler D, Luc G, Bard JM, Bara L, Ricard S. Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. *Nature* **359:**641–644, 1992.

36. Batzer M, Stoneking M, Alegria-Hartman M, Bazan H, Kass D, Shaikh T, Novick G, Ioannou PA. African origin of human-specific polymorphic Alu insertions. *Proc Natl Acad Sci USA* **91:**12288–12292, 1994.

37. Stoneking M, Fontius JJ, Clifford SL, Soodyall H, Arcot SS, Saha N, Jenkins T, Tahir MA, Deininger PL, Batzer MA. Alu insertion polymorphisms and human evolution: Evidence for a larger population size in Africa. *Genome Res* **7:**1061–1071, 1997.

38. Makalowski W, Mitchell GA, Labuda D. Alu sequences in the coding regions of mRNA: A source of protein variability. *Trends Genet* **10:**188–193, 1994.

39. Britten RJ. Mobile elements inserted in the distant past have taken on important functions. *Gene* **205:**177–182, 1997.

40. Tiret L, Rigat B, Visvikis S, Breda C, Corvol P, Cambien F, Soubrier F. Evidence, from combined segregation and linkage analysis, that a variant of the angiotensin I-converting enzyme (ACE) gene controls plasma ACE levels. *Am J Hum Genet* **51:**197–205, 1992.

41. Oldridge M, Zackai EH, McDonald-McGinn DM, Iseki S, Morriss-Kay GM, Twigg SR, Johnson D, Wall SA, Jiang W, Theda C, Jabs EW, Wilkie AO. De novo Alu-element insertions in FGFR2 identify a distinct pathological basis for Apert syndrome. *Am J Hum Genet* **64:**446–461, 1999.

42. Rudiger NS, Gregersen N, Kielland-Brandt MC. One short well conserved region of Alu-sequences is involved in human gene rearrangements and has homology with prokaryotic chi. *Nucleic Acids Res* **23:**256–260, 1995.

43. Lehrman MA, Russell DW, Goldstein JL, Brown MS. Alu-Alu recombination deletes splice acceptor sites and produces secreted low density lipoprotein receptor in a subject with familial hypercholesterolemia. *J Biol Chem* **262:**3354–3361, 1987.

44. Jeffs AR, Benjes SM, Smith TL, Sowerby SJ, Morris CM. The BCR gene recombines preferentially with Alu elements in complex BCR-ABL translocations of chronic myeloid leukaemia. *Hum Mol Genet* **7:**767–776, 1998.

45. Chen SJ, Chen Z, Font MP, d'Auriol L, Larsen CJ, Berger R. Structural alterations of the BCR and ABL genes in Ph1 positive acute leukemias with rearrangements in the BCR gene first intron: Further evidence implicating Alu sequences in the chromosome translocation. *Nucleic Acids Res* **17:**7631–7642, 1989.

46. Super HG, Strissel PL, Sobulo OM, Burian D, Reshmi SC, Roe B, Zeleznik-Le NJ, Diaz MO, Rowley JD. Identification of complex genomic breakpoint junctions in the t(9;11) MLL-AF9 fusion gene in acute leukemia. *Genes Chromosomes Cancer* **20:**185–195, 1997.

47. Hasty P, Rivera-Perez J, Bradley A. The length of homology required for gene targeting in embryonic stem cells. *Mol Cell Biol* **11:**5586–5591, 1991.

48. Burwinkel B, Kilimann MW. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *J Mol Biol* **277:**513–517, 1998.

49. Shen M, Deininger P. An *in vivo* assay for measuring the recombination potential between DNA sequences in mammalian cells. *Anal Biochem* **205:**83–89, 1992.

50. Kazazian HH, Jr. Mobile elements and disease. *Curr Opin Genet Dev* **8:**343–350, 1998.

51. Janicic N, Pausova Z, Cole DE, Hendy GN. Insertion of an Alu sequence in the Ca(2+)-sensing receptor gene in familial hypocalciuric hypercalcemia and neonatal severe hyperparathyroidism. *Am J Hum Genet* **56:**880–886, 1995.

52. Economou-Pachnis A, Tsichlis PN. Insertion of an Alu SINE in the human homologue of the Mlvi-2 locus. *Nucleic Acids Res* **13:**8379–8387, 1985.

53. Wallace MR, Andersen LB, Saulino AM, Gregory PE, Glover TW, Collins FS. A de novo Alu insertion results in neurofibromatosis type 1. *Nature* **353:**864–866, 1991.

54. Rowe SM, Coughlan SJ, McKenna NJ, Garrett E, Kieback DG, Carney DN, Headon DR. Ovarian carcinoma-associated TaqI restriction fragment length polymorphism in intron G of the progesterone receptor gene is due to an Alu sequence insertion. *Cancer Res* **55:**2743–2745, 1995.

55. Lester T, McMahon C, VanRegemorter N, Jones A, Genet S. X-linked immunodeficiency caused by insertion of Alu repeat sequences. *J Med Gen Suppl* **34**(Suppl 1): S81, 1997.

56. Vidaud D, Vidaud M, Bahnak BR, Siguret V, Gispert Sanchez S, Laurian Y, Meyer D, Goossens M, Lavergne JM. Haemophilia B due to a de novo insertion of a human-specific Alu subfamily member within the coding region of the factor IX gene. *Eur J Hum Genet* **1:**30–36, 1993.

57. Abdelhak S, Kalatzis V, Heilig R, Compain S, Samson D, Vincent C, Levi-Acobas F, Cruaud C, Le Merrer M, Mathieu M, Konig R, Vigneron J, Weissenbach J, Petit C, Weil D. Clustering of mutations responsible for branchio-oto-renal (BOR) syndrome in the eyes absent homologous region (eyaHR) of EYA1. *Hum Mol Genet* **6:**2247–2255, 1997.

58. Muratani K, Hada T, Yamamoto Y, Kaneko T, Shigeto Y, Ohue T, Furuyama J, Higashino K. Inactivation of the cholinesterase gene by Alu insertion: Possible mechanism for human gene transposition. *Proc Natl Acad Sci USA* **88:** 11315–11319, 1991.

59. Halling KC, Honchel R, Lazzaro CR, Bufill JA, Arndt C, Lindor NM. A germline Alu1 repeat insertion of the APC gene leading to hereditary desmoid disease in an Amish family. *Am J Hum Genet Suppl* **61:**A67, 1997.

60. Stoppa-Lyonnet D, Carter PE, Meo T, Tosi M. Clusters of intragenic Alu repeats predispose the human C1 inhibitor locus to deleterious rearrangements. *Proc Natl Acad Sci* **87:**1551–1555, 1990.

61. Miki Y, Katagiri T, Kasumi F, Yoshimoto T, Nakamura Y.

Mutation analysis in the BRCA2 gene in primary breast cancers. *Nature Genet* **13:**245–247, 1996.

62. Zhang Y-H, Huang B-L, Finlayson G, Deininger PL, McCabe ERB. Alu Sx insertion in a patient with benign glycerol kinase deficiency. *Am J Hum Genet Suppl* **63:**A395, 1998.
63. Chae JJ, Park YB, Kim SH, Hong SS, Song GJ, Han KH, Namkoong Y, Kim HS, Lee CC. Two partial deletion mutations involving the same Alu sequence within intron 8 of the LDL receptor gene in Korean patients with familial hypercholesterolemia. *Hum Genet* **99:**155–163, 1997.
64. Lehrman MA, Goldstein JL, Russel DW, Brown MS. Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell* **48:**827–835, 1987.
65. Lehrman MA, Schneider WJ, Sudhof TC, Brown MS, Goldstein JL, Russell DW. Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* **227:**140–146, 1985.
66. Rudiger NS, Heinsvig EM, Hansen FA, Faergeman O, Bolund L, Gregersen N. DNA deletions in the low density lipoprotein (LDL) receptor gene in Danish families with familial hypercholesterolemia. *Clin Genet* **39:**451–462, 1991.
67. Yamakawa K, Takada K, Yanagi H, Tsuchiya S, Kawai K, Nakagawa S, Kajiyama G, Hamaguchi H. Three novel partial deletions of the low-density lipoprotein (LDL) receptor gene in familial hypercholesterolemia. *Hum Genet* **82:**317–321, 1989.
68. Flint J, Rochette J, Craddock CF, Dode C, Vignes B, Horsley SW, Kearney L, Buckle VJ, Ayyub H, Higgs DR. Chromosomal stabilisation by a subtelomeric rearrangement involving two closely related Alu elements. *Hum Mol Genet* **5:**1163–1169, 1996.
69. Ko TM, Tseng LH, Kao CH, Lin YW, Hwa HL, Hsu PM, Li SF, Chuang SM. Molecular characterization and PCR diagnosis of Thailand deletion of alpha-globin gene cluster. *Am J Hematol* **57:**124–130, 1998.
70. Harteveld KL, Losekoot M, Fodde R, Giordano PC, Bernini LF. The involvement of Alu repeats in recombination events at the alpha-globin gene cluster: Characterization of two alphazero-thalassaemia deletion breakpoints. *Hum Genet* **99:**528–534, 1997.
71. Nicholls RD, Fischel-Ghodsian N, Higgs DR. Recombination at the human alpha-globin gene cluster: Sequence features and topological constraints. *Cell* **49:**369–378, 1987.
72. Ariga T, Carter PE, Davis AE. Recombinations between Alu repeat sequences that result in partial deletions within the C1 inhibitor gene. *Genomics* **8:**607–613, 1990.
73. Pousi B, Hautala T, Heikkinen J, Pajunen L, Kivirikko KI, Myllyla R. Alu-Alu recombination results in a duplication of seven exons in the lysyl hydroxylase gene in a patient with the type VI variant of Ehlers-Danlos syndrome. *Am J Hum Genet* **55:**899–906, 1994.
74. Hu XY, Ray PN, Worton RG. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J* **10:**2471–2477, 1991.
75. Markert ML, Hutton JJ, Wiginton DA, States JC, Kaufman RE. Adenosine deaminase (ADA) deficiency due to deletion of the ADA gene promoter and first exon by homologous recombination between two Alu elements. *J Clin Invest* **81:**1323–1327, 1988.
76. Huang LS, Ripps ME, Korman SH, Deckelbaum RJ, Breslow JL. Hypobetalipoproteinemia due to an apolipoprotein B gene exon 21 deletion derived by Alu-Alu recombination. *J Biol Chem* **264:**11394–11400, 1989.
77. Shimada F, Taira M, Suzuki Y, Hashimoto N, Nozaki O, Tatibana M, Ebina Y, Tawata M, Onaya T. Insulin-resistant diabetes associated with partial deletion of insulin-receptor gene. *Lancet* **335:**1179–1181, 1990.
78. Kornreich R, Bishop DF, Desnick RJ. Alpha-galactosidase A gene rearrangements causing Fabry disease. Identification of short direct repeats at breakpoints in an Alu-rich gene. *J Biol Chem* **265:**9319–9326, 1990.
79. Marcus S, Hellgren D, Lambert B, Fallstrom SP, Wahlstrom J. Duplication in the hypoxanthine phosphoribosyl-transferase gene caused by Alu-Alu recombination in a patient with Lesch Nyhan syndrome. *Hum Genet* **90:**477–482, 1993.
80. Li L, Bray PF. Homologous recombination among three intragene Alu sequences causes an inversion-deletion resulting in the hereditary bleeding disorder Glanzmann thrombasthenia. *Am J Hum Genet* **53:**140–149, 1993.
81. Burwinkel B, Kilimann MW. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *J Mol Biol* **277:**513–517, 1998.
82. Hori T, Tomatsu S, Nakashima Y, Uchiyama A, Fukuda S, Sukegawa K, Shimozawa N, Suzuki Y, Kondo N, Horiuchi T. Mucopolysaccharidosis type IVA: Common double deletion in the *N*-acetylgalactosamine-6-sulfatase gene (GALNS). *Genomics* **26:**535–542, 1995.
83. Olds RJ, Lane DA, Chowdhury V, De SV, Leone G, Thein SL. Complete nucleotide sequence of the antithrombin gene: Evidence for homologous recombination causing thrombophilia. *Biochemistry* **32:**4216–4224, 1993.
84. Rouyer F, Simmler MC, Page DC, Weissenbach J. A sex chromosome rearrangement in a human XX male caused by Alu-Alu recombination. *Cell* **51:**417–425, 1987.
85. Myerowitz R, Hogikyan ND. A deletion involving Alu sequences in the beta-hexosaminidase alpha-chain gene of French Canadians with Tay-Sachs disease. *J Biol Chem* **262:**15396–15399, 1987.
86. Botto M, Fong KY, So AK, Barlow R, Routier R, Morley BJ, Walport MJ. Homozygous hereditary C3 deficiency due to a partial gene deletion. *Proc Natl Acad Sci USA* **89:**4957–4961, 1992.
87. Neote K, McInnes B, Mahuran DJ, Gravel RA. Structure and distribution of an Alu-type deletion mutation in Sandhoff disease. *J Clin Invest* **86:**1524–1531, 1990.
88. Strout MP, Marcucci G, Bloomfield CD, Caligiuri MA. The partial tandem duplication of ALL1 (MLL) is consistently generated by Alu-mediated homologous recombination in acute myeloid leukemia. *Proc Natl Acad Sci USA* **95:**2390–2395, 1998.
89. So CW, Ma ZG, Price CM, Dong S, Chen SJ, Gu LJ, So CK, Wiedemann LM, Chan LC. MLL self fusion mediated by Alu repeat homologous recombination and prognosis of AML-M4/M5 subtypes. *Cancer Res* **57:**117–122, 1997.
90. Schichman SA, Caligiuri MA, Strout MP, Carter SL, Gu Y,

Canaani E, Bloomfield CD, Croce CM. ALL-1 tandem duplication in acute myeloid leukemia with a normal karyotype involves homologous recombination between Alu elements. *Cancer Res* **54:**4277–4280, 1994.

91. Swensen J, Hoffman M, Skolnick MH, Neuhausen SL. Identification of a 14 kb deletion involving the promoter region of BRCA1 in a breast cancer family. *Hum Mol Genet* **6:**1513–1517, 1997.
92. Puget N, Sinilnikova O, Stoppa-Lyonnet D, Ayyadevera R, Pages S, Lynch H, Goldgar D, Lenoir GM, Mazoyer S. An Alu-mediated 6-kb duplication in the BRCA1 gene: A new founder mutation? *Am J Hum Genet* **64:**300–302, 1999.
93. Mauillon JL, Michel P, Limacher JM, Latouche JB, Dechelotte P, Charbonnier F, Martin C, Moreau V, Metayer J, Paillot B, Frebourg T. Identification of novel germline hMLH1 mutations including a 22 kb Alu-mediated deletion in patients with familial colorectal cancer. *Cancer Res* **56:** 5728–5733, 1996.
94. Onno M, Nakamura T, Hillova J, Hill M. Rearrangement of the human tre oncogene by homologous recombination between Alu repeats of nucleotide sequences from two different chromosomes. *Oncogene* **7:**2519–2523, 1992.
95. Rothberg PG, Ponnuru S, Baker D, Bradley JF, Freeman AI, Cibis GW, Harris DJ, Heruth DP. A deletion polymorphism due to Alu-Alu recombination in intron 2 of the retinoblastoma gene: Association with human gliomas. *Mol Carcinog* **19:**69–73, 1997.
96. Zucman-Rossi J, Batzer MA, Stoneking M, Delattre O, Thomas G. Interethnic polymorphism of EWS intron 6: Genome plasticity mediated by Alu retroposition and recombination. *Hum Genet* **99:**357–363, 1997.