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

Naoko Narita, Hisahide Nishio, Yoshihiko Kitoh, Yuka Ishikawa,‡ Yukitoshi Ishikawa,‡ Ryoji Minami,‡ Hajime Nakamura, and Masafumi Matsuo*

*Department of Pediatrics and *International Center for Medical Research, Kobe University School of Medicine, Kobe 650 Japan; and ‡Department of Pediatrics, National Sanatorium Yakumo Hospital, Hokkaido 049-31, Japan*

Abstract

We report here the second evidence of retrotransposition of L1, which was found inserted into the dystrophin gene of a patient, causing Duchenne muscular dystrophy (DMD). When the PCR was used to amplify a region of the dystrophin gene encompassing exon 44 from genomic DNA of two Japanese brothers with DMD, it was found to be ~ 600 bp larger than expected. Both the normal and the abnormally large products were amplified from the DNA of their mother. However, the maternal grandparents did not have the abnormal allele, and the mutation must therefore have occurred in the mother. Analysis of nucleotide sequence of the amplified product from a patient disclosed that the insertion was present zero to two bases upstream from the 3′ end of exon 44 and that two to four bases of the exon sequence were deleted from the insertion site. The insertion sequence was found to be composed of 606–608 bp and to be almost identical to the inverse complement of 3′ portion of the L1 retrotransposon consensus sequence. The dystrophin gene transcript from peripheral lymphocytes of one of the patients was analyzed by using reverse transcription/semi-nested PCR. The size of the amplified product encompassing exon 42 to 46 was smaller than expected. Sequencing of the amplified product disclosed that the sequence of exon 43 was directly joined to that of exon 45. Exon 44 of the transcript was thus shown to be skipped during splicing. This novel mutation of the dystrophin gene has important implications regarding retrotransposition of an active L1 element and provides a new insight into the origins of mutations in the dystrophin gene. (*J. Clin. Invest.* 1993. 91:1862–1867.) Key words: Duchenne muscular dystrophy • dystrophin gene • exon skipping • L1 • retrotransposition

Introduction

Duchenne and Becker muscular dystrophies (DMD[1] and BMD, respectively) are caused by mutations in the dystrophin gene, which is spread over 2,500 kb of the X chromosome. A large proportion of mutations are partial gene deletions or duplications (60 and 6–10% of the total, respectively) of varying sizes (1–3). Results of gene analyses show that in > 90% of cases, mutations that maintain the reading frame in the dystrophin transcript cause BMD, while those in which a frameshift has occurred cause DMD (4). Recently, we found that determination of translational reading frame by analyzing dystrophin mRNA is far more reliable than gene analysis for the diagnosis of DMD/BMD because exon skipping during splicing resulting from deletion of a very small segment of exon sequence has been demonstrated to occur in dystrophin Kobe (5).

DMD is frequent (1 per 3,500 live male births), and it has been estimated that one third of DMD/BMD cases result from new mutations of the dystrophin gene (6). However, other than its large size, there is no obvious explanation for the apparently abnormally high rate of mutations in the dystrophin gene. The dystrophin gene has been shown to have two mutation hot spots, but they appear to be devoid of features that might render them susceptible to mutation (1).

L1 elements are a family of long, interspersed, repetitive DNA sequences that are present at ~ 10^5 copies (full length and truncated) dispersed throughout the genome. The human full-length L1 element comprises 6.1 kb, and contains two long open reading frames (ORFs), the second of which could encode a polypeptide having sequence similarity to reverse transcriptases, and a polyadenylation signal followed by poly (A) tract. The majority of L1 elements are truncated at their 5′ end and have premature stop codons in the ORF. Most of the elements have flanking target site duplications of 5 to 15 bases. These structural characters suggest that L1 elements are capable of transposition by way of an RNA intermediate (7). Evidence that some L1 elements in the human genome are retrotransposable was provided by de novo transposition of a truncated L1 DNA into exon 14 of the Factor VIII gene on the X chromosome in two individuals (8). Each new insertion contained the 3′ portion of an L1 element, including much of the second open reading frame of the consensus sequence and a poly(A) tract, and was flanked by a perfect target site duplication. From its unique 3′ trailing sequence, the inserted sequences have been assigned to a subtype of L1 family (9). In addition, a full-length L1 element considered to be a progenitor of the inserted sequence was identified on chromosome 22 (10) and shown to have retrotranscriptase activity (11).

In this paper, we report that a 5′ truncated L1 element was inserted into an exon sequence of the dystrophin gene in two Japanese brothers with DMD, and that the mutation, which originated in their mother, resulted in exon skipping during splicing of the dystrophin mRNA precursor.

Address correspondence to Masafumi Matsuo, M.D., Ph.D., International Center for Medical Research, Kobe University School of Medicine, 7-5-1 Kusunokicho, Chuo, Kobe 650, Japan.

Received for publication 31 August 1992 and in revised form 21 December 1992.

1. *Abbreviations used in this paper:* BMD, Becker muscular dystrophy; DMD, Duchenne muscular dystrophy; ORF, open reading frame.

J. Clin. Invest.
© The American Society for Clinical Investigation, Inc.
0021-9738/93/05/1862/06 $2.00
Volume 91, May 1993, 1862–1867

Methods

Cases. A pair of Japanese brothers (20 and 19 yr old) born in the northern part of Japan were enrolled in this study. They were diagnosed as having DMD based on the findings of high serum creatine kinase, electromyogram, and muscle biopsy. The elder and the younger brothers became wheelchair-bound at the ages of 11 and 10, respectively, and were admitted to National Sanatorium Yakumo Hospital. They now need respiratory support from an artificial ventilator. Their mother was diagnosed as a carrier of DMD on the basis of high serum creatine kinase activity but showed no signs of muscle weakness. Maternal grandparents had neither high serum creatine kinase activity nor signs of muscle weakness. Informed consent was obtained to perform molecular genetic studies on the family.

Amplification of genomic DNA. For the screening of deletion mutations of the dystrophin gene, exon-containing fragments were amplified by using primers corresponding to regions in flanking introns (12, 13). Extraction of DNA from blood cells and the PCR were carried out according to methods previously described (14). An aliquot of amplified DNA was electrophoresed on a 3% agarose gel and stained with ethidium bromide along with low molecular size DNA standards. The position of primers used to amplify the exon 44–containing region are shown in Fig. 1 *a*. As the restriction enzyme recognition site for AflII is present at the 3′ end of exon 44, the amplified product was digested with this enzyme according to the method suggested by the supplier (New England Biolabs, Beverly, MA) and the digested fragments were electrophoresed in a 3% agarose gel. To localize roughly the insertion site, primers based both on the exon 44 sequence and complementary to the exon 44 sequence (Fig. 1 *a*) were synthesized on a DNA synthesizer (Applied Biosystems, Inc., Foster City, CA) and the exon 44–containing region was amplified as two separate fragments.

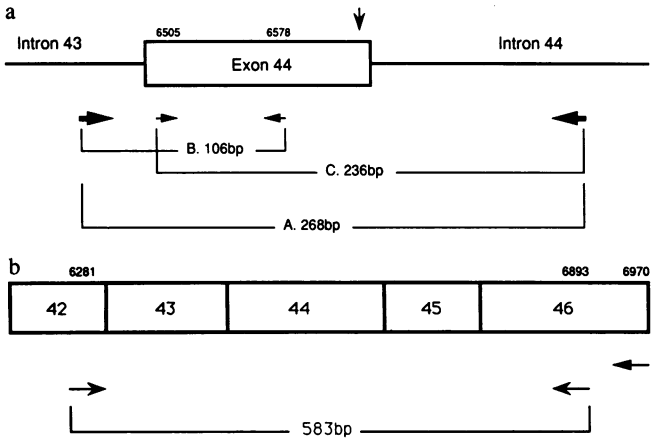

Figure 1. Schematic presentation of amplified region. (*a*) Amplification of genomic DNA. The box represents exon 44 and the lines represent introns 43 and 44. Horizontal arrows indicate the positions and the orientations of the primers. Numbers over the exon indicate nucleotides in the dystrophin cDNA (36) where the 5′ end of primers are located. Brackets indicate the size of amplified products. Vertical arrow shows the recognition site for AflII. For deletion screening, primers indicated by bold arrows on introns were used to amplify, and the size of the amplified product is calculated to be 268 bp (bracket *A*). To amplify the fragment as two separate parts, primers (*bold arrows*) were paired with each corresponding primers (*thin arrows*) (brackets *B* and *C*). The amplified product shown by brackets *B* and *C* are 106 and 236 bp long, respectively. The sequences of the oligonucleotide primers (*bold arrows*) used for deletion screening have been described elsewhere (12). (*b*) Amplification of dystrophin cDNA. Three primers were designed for semi-nested PCR. For the first amplification, a larger fragment was amplified using outer set of primers, one based on the sequence of exon 42 (5′-GCTAAGGACTTTGAAGATCT-3′) and the other complementary to a sequence in exon 46 (5′-CTTGACTTGCTCAAGCTTTTC-3′). For the second amplification, an inner reverse primer complementary to a sequence of exon 46 (5′-CTCTTTGAAATTCTGACAAG-3′) was used. Bracket indicates the size of amplified products which is calculated as 583 bp. Numbers in boxes indicate exon numbers and numbers over the exon indicate nucleotides in the dystrophin cDNA where the 5′ end of primers are located.

Amplification of dystrophin cDNA in lymphocytes. Reverse transcription/semi-nested PCR was used to amplify the region across the junctions of exons 42, 43, 44, 45, and 46 of dystrophin mRNA in lymphocytes. Total RNA was extracted from peripheral lymphocytes (1×10^6 cells) by the method of Chomczynski et al. (15) and cDNA was prepared as described before (5). Primers (20–21 bases long) designed to amplify this region were synthesized. The position of primers used are shown in Fig. 1 *b*. The 583-bp long target region extended from the last 45 bp of exon sequence 42 to the first 45 bp of exon 46 was expected to be amplified after semi-nested PCR. PCR amplification of 35 cycles with the outer pair of primers (5) was followed by a second PCR amplification using the inner reverse primer. For this second amplification of 30 cycles, 1 μl of the 20 μl of the first amplification mixture was transferred to the second amplification mixture.

Sequencing of PCR products. The amplified DNA was treated with T4 DNA polymerase (Life Technologies, Inc., Grand Island, NY) and subcloned into SmaI-digested PGEM 7Zf(+) (Promega Co., Madison, WI). The sequences of inserted DNA from three or four clones were determined by an automatic DNA sequencer (model 373A; Applied Biosystems Inc.) using the Taq dye primer cycle sequencing kit (Applied Biosystems Inc.).

Results

We have previously screened for deletion mutations in the dystrophin gene by using PCR (13, 16). Nine exons distributed throughout the deletion-prone region of the dystrophin gene have been amplified from genomic DNA from over 200 Japanese patients with DMD/BMD. In the cases discussed here, agarose gel electrophoresis did not reveal the presence of an amplified product of the expected 268 bp fragment encompassing exon 44, which was apparently replaced by a larger amplified product (Fig. 2, lanes *4* and *5*). All other amplified exons were of the normal size. These results indicated the presence of an insertion mutation in the amplified region. To determine the location of the insertion more precisely, the oversized frag-

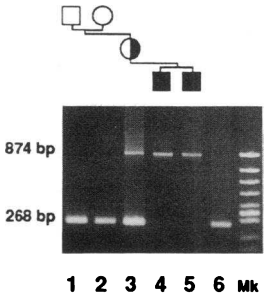

Figure 2. Amplification of exon 44 encompassing region of the dystrophin gene from family members. The fragments amplified from patients migrated more slowly (lanes *4* and *5*) than that from a normal individual (lane *6*). From its mobility, the length of the patient's amplified fragment was calculated to be ~ 900 bp, implying roughly a 600-bp increment compared to the normal fragment (268 bp). Segregation of the insertion was studied in the family members. The mother (lane *3*) had both normal and larger amplified products and is therefore a carrier. Only the normal amplified product was obtained from the maternal grandparents (lanes *1* and *2*). (*Top*) Pedigree with squares representing males and circles representing females. The affected individual is represented by a filled symbol and carriers are represented by a half-filled symbol. Mk refers to HincII digested φX 174 phage DNA (Toyobo Co., Osaka, Japan).

ment was amplified as two separate parts using primers both complementary to central region of exon 44 sequence and based on 5′ end of exon 44 (Fig. 1 *a*). The amplified product encompassing intron 43 to the middle of exon 44 was as large as expected (106 bp), while the amplified product encompassing the 5′ end of exon 44 to intron 44 was ~ 600 bp larger than the normal product (~ 850 bp vs 236 bp) (data not shown). These results indicated that an insertion was present in the boundary region of exon 44 and intron 44.

To determine the inheritance of this mutation, the exon 44–encompassing fragment was also amplified from genomic DNA obtained from the mother and from the maternal grandparents. Two different products were amplified from the mother's DNA: one corresponding to the normal product and the other corresponding to the product amplified from the DNA of her sons (Fig. 2, lane *3*). This indicated that the mother is a carrier of this mutation. Only the normal-sized fragment was amplified from DNA from the maternal grandparents (Fig. 2, lanes *1* and *2*). All amplified products corresponding to normal fragments could be cut into two segments by restriction enzyme AflII (data not shown). These results indicated that the maternal grandparents had only normal allele and that the insertion mutation originated in the mother.

To characterize the insertion mutation further, the larger fragment amplified from the elder brother was sequenced. The nucleotide sequence of the 5′ portion of the fragment was completely homologous to that of the boundary region between intron 43 and exon 44 of the dystrophin gene up to nucleotide 6642 of dystrophin cDNA (Fig. 3). The region immediately downstream from nucleotide 6642 was completely different from that of the wild-type sequence, and contained a 606-bp long insertion starting with a C followed by 18 successive T nucleotides. Sequences homologous to the last two nucleotides of 3′ end of exon 44 sequence (nucleotides 6645 and 6646 of dystrophin cDNA) and the sequence of intron 44 were identified after the end of the insertion sequence (Fig. 3). Thus, the insertion sequence is 2 bp upstream from the 3′ end of exon 44, and two nucleotides of the exon sequence (TA, numbered 6643 and 6644 in the sequence of the dystrophin cDNA, respectively) appear to have been deleted. The insertion occurred within an A + T rich region in the dystrophin gene at the AflII restriction enzyme recognition site (CTTAAG: nucleotides 6641–6646 in the dystrophin cDNA) in the dystrophin gene. The mutation was named dystrophin Yakumo.

By screening DNA sequence data banks, the inserted sequence was found to be 98% identical to the inverse complement of the 3′ portion (nucleotides 5563 to 6161) of the consensus genomic sequence of the L1 element (named L1.HS) (17) (Fig. 4). Although some nucleotides of the insertion in dystrophin Yakumo (L1.DY) differed from the consensus sequence, the 3′ portion of the second ORF was conserved. L1.DY has a poly (A) tract preceded by a polyadenylation signal (AATAAA), but has an additional seven successive A nucleotides and one G nucleotide downstream of the 3′ end of the region that is highly homologous to L1.HS (Fig. 4). Among the 598 nucleotides that are homologous to the consensus L1 sequence, nine nucleotides of L1.DY differ from those in L1.HS. However, eight of these nine nucleotides are same as those in consensus L1 cDNA sequence (18). The sole difference between L1.DY and this cDNA consensus is the A to G transition at nucleotide 5823 (Fig. 4). L1 elements can be subtyped according to the sequence of the 3′ trailing region, and a consensus sequence has also been proposed for this region from the analysis of 20 randomly selected human genomic L1 elements (9). L1.DY has exactly same nucleotide sequence as the cDNA of the Ta subset of L1 elements (18). All of these data strongly suggest that retrotranscriptional insertion of an active L1 occurred very recently in the dystrophin gene.

As the insertion sequence in dystrophin Yakumo was found to be a 5′ truncated L1 element, and the two nucleotides at the 3′ end of the inserted element are the same as those at the 5′ end of the target site of exon 44 (Fig. 3), the insertion sequence might be two bases longer (608 bp), and four nucleotides of the 3′ end of exon 44 might have disappeared from the target site. In this extreme case, none of the exon 44 sequence is left at the 5′ end of disrupted dystrophin gene.

It is very important to determine the effect of mutations on the translational reading frame for the diagnosis of DMD/BMD. For this purpose, an illegitimate transcript of the dystrophin gene in peripheral lymphocytes has been used to determine the reading frame of dystrophin transcript in muscle cells (19). In this study, we analyzed the dystrophin transcript around exon 44 region by amplifying a region encompassing exons 42–46 on dystrophin cDNA (Fig. 1 *b*). A 583 bp long product was amplified from the normal transcript by semi-nested PCR (Fig. 5, lane *1*), while an ~ 600-bp longer fragment was expected to be amplified from cDNA prepared from lymphocytes of the elder brother. Instead, the amplified prod-

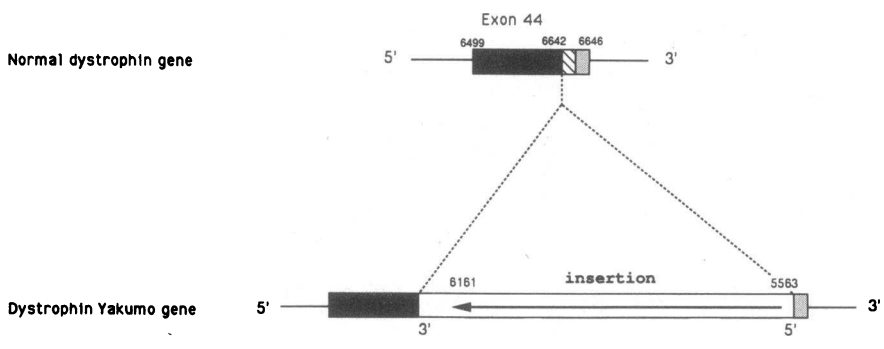

Figure 3. Scheme of insertion mutation into exon 44 of the dystrophin gene. The insertion sequence was present at a maximum of two bases upstream from 3′ end of exon 44 of the dystrophin gene and nucleotides T and A (numbered 6643 and 6644, respectively, and part of the AflII restriction enzyme cutting site) are missing. Three potential sites at which the insertion took place were proposed because of the common nucleotide sequence (AG) at the 3′ end of the insertion and the 5′ end of the disrupted gene. Filled and open boxes represent exon 44 and the L1 element, respectively. The arrow points towards the 3′ end of the L1 sequence. The box with diagonal lines corresponds to TA nucleotides which disappeared from the dystrophin gene of dystrophin Yakumo. The shaded box indicates common AG nucleotides sequence to the insertion and exon 44. Numbers over upper and lower scheme indicate positions in dystrophin cDNA (36) and the consensus nucleotide sequence of L1 (18), respectively.

```
GGGACTGTAAACTAGTTCAACCATTGTGGAAGTCAGTGTGGCGATTCCTC  5600   L1.HS
CAAAGACTTAC--------------------------------------    570   L1.DY

AGGGATCTAGAACTAGAAATACCATTTGACCCAGCAATCCCATTACTGGG  5650   L1.HS
--------------------------------------------------  520   L1.DY

TATATACCCAAAGGACTATAAATCATGCTGCTATAAAGACACATGCACAC  5700   L1.HS
----------T--G-------------------------------------  470   L1.DY

GTATGTTTATTGCGGCATTATTCACAATAGCAAAGACTTGGAACCAACCC  5750   L1.HS
--------------------------------------------------  420   L1.DY

AAATGTCCAACAATGATAGACTGGATAAAGAAAATGTGGCACATATACAC  5800   L1.HS
--------------------------------------------------  370   L1.DY

CATGGAATACTATGCAGCCATAAAAAAATGATGAGTTCATGTCCTTTGTAG 5850   L1.HS
----------------------G----------------A----------  320   L1.DY

GGACATGGATGAAATTGGAAACCATCATTCTTAGTAAACTATCGCAAGAA  5900   L1.HS
--------------------------------------------------  270   L1.DY
                                            ***
CAAAAAACCAAACACCGCATATTCTCACTCATAGGTGGGAATTGAACAAT  5950   L1.HS
--------------------------------------------------  220   L1.DY

GAGATCACATGGACACAGGAAGGGGAATATCACACTCTGGGGACTGTCGT  6000   L1.HS
-----------------------------------------------G--  170   L1.DY

GGGGTAGGGGGAGGGGGGAGGGATAGCATTGGGAGATATACCTAATGCTA  6050   L1.HS
-----C---------------------------------------------  120   L1.DY

GATGACAAGTTAGTGGGTGCAGCACACCAGCATGGCACATGTATACATAT  6100   L1.HS
-------CA------------------------------------------   70   L1.DY

GTAACTAACCTGCACAATGTGCACATGTACCCTAAAACTTAAAGTATAAAT 6150   L1.HS
--------------------------------------G-------====   20   L1.DY
          6161
AAAAAAAAAAA                                              L1.HS
==========AAAAAAAGAGATACCATTTGTATTTAGCATGTTCCCAAT]  -30   L1.DY
```

Figure 4. Comparison of the inserted sequences and L1 sequences. The inserted sequence (L1.DY) is shown below and numbered 5′ to 3′. The L1 consensus sequence (L1.HS) is shown above and numbered as reported (18). Boxes enclose the sequence of exon 44 of the dystrophin gene. On the insertion sequence dashes indicate nucleotides that are the same as in the L1 consensus sequence. The L1 insertion showed 98% homology to the consensus genomic L1 sequence outside of the 3′-trailer region. The nucleotides that differed from the L1 consensus sequence but matched those of the consensus L1 cDNA are shown. The underlined nucleotide is that which differs from the consensus L1 cDNA (A to G transition). The insertion is not flanked by a target site duplication. The bracket indicates the polyadenylation signal and asterisks correspond to stop codon.

uct was found to be ~ 150 bp smaller than normal (Fig. 5, lane *2*), which suggested that some aberrant splicing occurred in dystrophin Yakumo.

To study this in more detail, the abnormal amplified product from cDNA was sequenced. The sequence was completely devoid of exon 44, and the 3′ end of exon 43 was joined directly to the 5′ end of exon 45 (Fig. 6). This indicated that exon 44 of the patient's dystrophin gene was skipped during splicing of the mRNA precursor. The translational reading frame of the dystrophin gene was thereby shifted by 2 bases and a stop codon appeared at codon 2182 (Fig. 6). Considering that patients were suffering from DMD, this truncated dystrophin transcript was probably the main in situ spliced product in muscle cells.

Discussion

Evidence for the contemporary movement of an L1 element in humans was first obtained in cases of hemophilia A that resulted from de novo insertion of an L1 element into Factor VIII gene (8). The structural characteristics of L1 elements

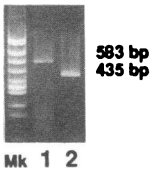

Figure 5. Amplification of cDNA prepared from lymphocyte mRNA. The amplified product encompassing exons 42 to 46 was obtained by semi-nested PCR. The expected fragment corresponding to 583 bp was obtained from the control (lane *1*). In contrast, neither the normal size fragment nor the fragment corresponding to insertion mutation were amplified from patient's cDNA (lane *2*). Instead, the amplified fragment from the patient migrated faster than that from the normal individual. From its mobility, the length of the patient's amplified fragment was calculated to be ~ 435 bp, implying roughly a 150-bp reduction compared to the wild-type fragment. Mk refers to HincII digested ϕX 174 phage DNA (Toyobo Co.).

indicate that they can amplify themselves by a cycle of transcription, reverse transcription, and integration in germline cells or in cells designated to become germlines. By using the largest L1 element inserted into Factor VIII gene as a probe, Dombroski et al. were able to find a full-length L1 element (named L1.2) that is considered to be a progenitor of the L1 element inserted into Factor VIII gene, and to map it on chromosome 22 (10). Furthermore, it was shown that the protein encoded by L1.2 has reverse transcriptase activity (11). These findings support a self-propagation model for the spread of the L1 element through the human genome (10).

Dystrophin Yakumo described here is the second example of the disruption of a human gene induced by the insertion of an L1 element. The mutation is considered to be a fairly rare and relatively recent event, since the insertion of an L1 element into the dystrophin gene was found only in this family among > 200 Japanese families of dystrophy patients screened. The insertion event was found to have occurred in the mother (Fig. 2). The L1 element found in dystrophin Yakumo (L1.DY) was 98% homologous to 5′ truncated consensus L1 element (Fig. 4). The inserted sequence has the following characteristics: (*a*) The insertion occurred in an A + T rich region of the dystrophin gene. (*b*) L1.DY retains the ORF for retrotranscriptase, although it is truncated at the 5′ end, and has a polyadenylation signal at its 3′ end followed by poly (A) tract (Fig. 4). (*c*) L1.DY differs by only one nucleotide from the consensus sequence of a subtype of L1 which is known to be active retroposon (10).

In dystrophin Yakumo, the poly (T) tail of L1 cDNA may be involved in base-pairing with the A-rich dystrophin sequence after the introduction of staggered single-strand breaks. This break may have occurred at the palindromic structure that includes the AflII restriction enzyme recognition site. However, sequencing of genome showed that gap repair occurred, causing the loss of least two nucleotides of dystrophin sequence (Fig. 3). In addition, one G nucleotide was detected at the end of poly (A) tract (Fig. 4). A guanosine is known to be present in this position in the probable progenitor of L1.2 (10). Therefore, the last G nucleotide in L1.DY may be transcribed from the flanking sequence of the progenitor element. This suggests that the poly (A) tract was not added after transcription, but is rather transcribed from the genomic element.

One of the most intriguing findings reported here is that the insertion of L1.DY did not create a target site duplication (Fig. 4), which is considered to be one of the hallmarks of retroposons. However, the absence of target site duplications has been described for some other L1 elements (20).

The results of the study presented here suggest that at least

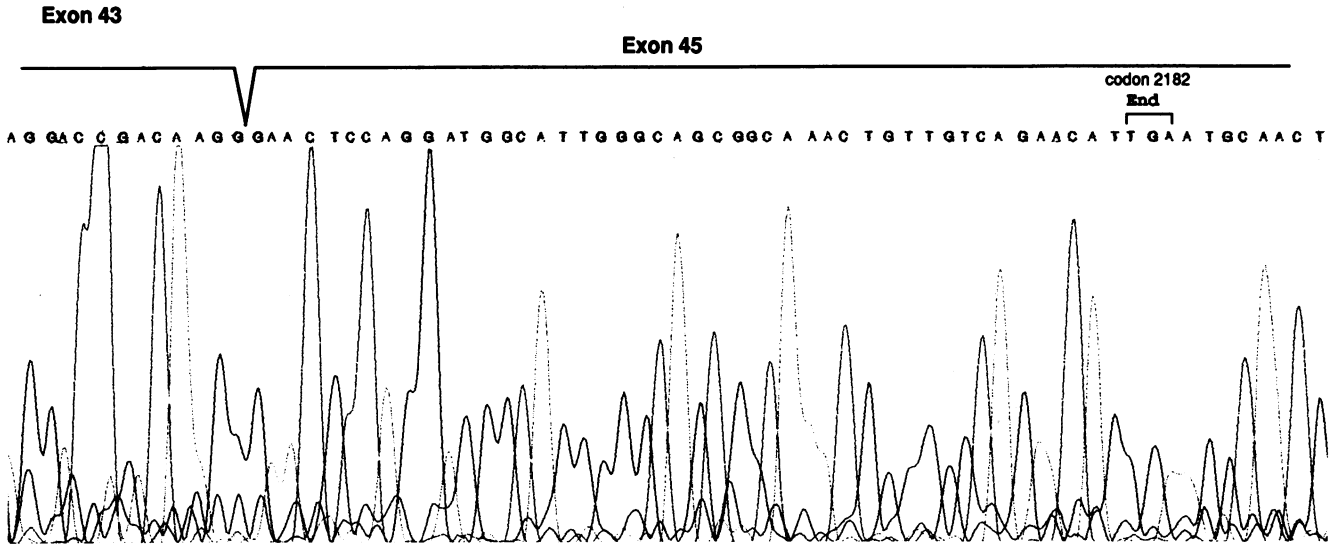

Figure 6. Sequence of the amplified product from patient's cDNA. Sequences of the region joining exons 43 to 45 are presented. The sequence of the 3′ end of exon 43 (5′-ACAAGG-3′) is joined directly to the 5′ end of exon 45 (5′-GAACTCC-3′). Because of this exon skipping, a stop codon (TGA) appears at codon 2182.

some copies of the L1 elements in the human genome are functionally active. On the basis of nucleotide sequences of their 3′ trailing regions, L1 elements have been divided into several subtypes (9). The active L1 element L1.2 (10) has almost same sequence in the 3′ trailing region as the consensus sequence of L1 cDNA. The 3′ trailing region of the L1.DY element is completely homologous to that of L1.2. This strongly suggests that L1.DY also originated from the same progenitor sequence as that which gave rise to the sequences inserted into Factor VIII gene (8).

Differing from germline cells, the L1 element has been reported to be functionally active in neoplastic cells (21). Brathauer et al. suggested that retrotransposition of the L1 element may be involved in the neoplastic process (22). Recently Miki et al. reported that a truncated L1 element has been inserted into exon sequence of the APC gene, a tumor suppressor gene, of colon cancer cells (23).

Exon skipping during splicing of mRNA precursors has been reported in some genomic mutations. The primary effect of almost every mutation which induces exon skipping during splicing is to change the consensus splicing sequence. This is not the case with dystrophin Kobe, however, which has a deletion mutation near the 3′ end of exon 19. We have shown that exon 19 of the dystrophin transcript is skipped during splicing even though the consensus splicing sequence is unaffected (5). In dystrophin Yakumo, exon 44 of the dystrophin transcript was found to be skipped during splicing (Fig. 6), indicating that the original 5′ splicing donor site of intron 44 of the dystrophin gene was inactivated because of the L1 insertion in the 3′ end of exon 44. In one-third of mutations with 5′ splice site inactivation, activation of cryptic splice sites has been reported (24). The insertion mutation of dystrophin Yakumo created a sequence of CAGGTTAGT, which had only one nucleotide difference from the consensus sequence (CAGGTAAGT) (25), at 57 bp downstream of the insertion site. Though this was supposed to be a candidate sequence for splice donor site, no amplified product corresponding to this splice site was obtained from dystrophin cDNA prepared from lymphocytes (Fig. 5, lane *2*). This suggest that factors other than a nucleotide sequence may determine the splice donor site (26, 27).

In dystrophin Yakumo, the consensus sequence for the 5′ splice donor site has been displaced to ~ 600 bp downstream from its original site due to the insertion mutation. As a result, an A nucleotide located 5′ end of the consensus splicing sequence was replaced by a T, but this change is not considered to be critical for splice site recognition (25). According to the model for exon recognition, the size of exon is at most 300 bp (27); the length of exon 44 with the insertion mutation (754 bp) is considerably larger. Therefore, the exon may not be recognized by the splicing machinery and therefore is skipped. Although the dystrophin transcript in lymphocytes is considered to be spliced in the same way as in muscle cells, further studies are needed to examine splicing in muscle cells from the dystrophin Yakumo. Unfortunately, this cannot be done because most of the patients' muscle have been replaced by connective and fatty tissues. Furthermore, we need to determine whether cryptic splice site activation occurs as a result of this mutation, as has been reported in some other exon skipping cases (28, 29).

By using in situ hybridization, the L1 family of retrotransposons was found to be present dominantly in Giemsa/Quinacrine-positive bands of chromosomes which are rich in A + T residues (30). L1 elements have been reported to be present in intron sequences of the dystrophin gene approximately every 30 kb (31). In the present paper we report that the dystrophin gene is a target for L1 insertion. From these findings, we hypothesize that unequal crossing over between two related L1 elements, which results in deletion or tandemly duplication of a region of the DNA sequence, might be a major cause of mutations of the dystrophin gene. This hypothesis is supported by the following evidence: (*a*) More than 60% of mutations of the dystrophin gene are deletions or duplications. (*b*) A small segment of an L1 sequence has been found between the breakpoints of a deletion in the β-globin gene cluster (32). (*c*) Fitch et al. showed that duplication of the γ-globin gene resulted from unequal homologous crossover between two related L1

elements flanking the single ancestral γ globin gene (33). (*d*) Two deletion breakpoints in the dystrophin gene which have been sequenced are bordered by nucleotide sequences highly homologous to the L1 consensus sequence (34). (*e*) Homologous recombination between Alu sequences, the other repetitive sequence of the human genome, have been shown to result in the deletion of a receptor gene coding for low density lipoprotein (35). Further studies are needed to test this hypothesis for the origin of mutations in the dystrophin gene.

Acknowledgments

We thank Dr. A. Pugsley for advice and for critically reading the manuscript.

This work was supported by grants from the Ministry of Education, Science and Culture of Japan (Nos. 04272212 and 04670596), the National Center of Neurology and Psychiatry of the Ministry of Health and Welfare of Japan, and the Morinaga Hoshikai Foundation.

References

1. Koenig, M., E. P. Hoffman, C. Bertelson, A. P. Monaco, C. Feener, and L. M. Kunkel. 1987. Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell.* 50:509–517.

2. Den Dunnen, J. T., P. M. Grootscholten, E. Bakker, L. A. J. Blonden, H. B. Ginjaar, M. C. Wapennar, H. M. B. Van Paassen, C. Van Broeckhoven, P. L. Pearson, and G. J. B. Van Ommen. 1989. Topography of the Duchenne muscular dystrophy (DMD) gene: FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Am. J. Hum. Genet.* 45:835–847.

3. Hu, X., P. N. Ray, E. G. Murphy, M. W. Thompson, and R. G. Worton. 1990. Duplicational mutation at the Duchenne muscular dystrophy locus: its frequency, distribution, origin, and phenotype genotype correlation. *Am. J. Hum. Genet.* 46:682–695.

4. Monaco, A. P., C. J. Bertelson, S. Liechti-Gallati, H. Moser, and L. M. Kunkel. 1988. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics.* 2:90–95.

5. Matsuo, M., T. Masumura, H. Nishio, T. Nakajima, Y. Kitoh, T. Takumi, J. Koga, and H. Nakamura. 1991. Exon skipping during splicing of dystrophin mRNA precursor due to an intra-exon deletion in the dystrophin gene of Duchenne muscular dystrophy Kobe. *J. Clin. Invest.* 87:2127–2131.

6. Moser, H. 1984. Duchenne muscular dystrophy: pathogenetic aspects and genetic prevention. *Hum. Genet.* 66:17–41.

7. Hutchinson, C. A., S. C. Hardies, D. D. Loeb, W. R. Shehee, and M. H. Edgell. 1989. LINEs and related retroposons: long interspersed repeated sequences in the eukaryotic genome. *In* Berg D, Howe M, ed. Mobile DNA. D. Berg and M. Howe, editors. American Society for Microbiology, Washington, DC. 593–617.

8. Kazazian, H. H., C. Wong, H. Youssoufian, A. F. Scott, D. G. Phillips, and S. E. Antonarakis. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature (Lond.).* 332:164–166.

9. Woods-Samuels, P., C. Wong, S. L. Mathias, A. F. Scott, H. H. Kazazian, and S. E. Antonarrakis. 1989. Characterization of a nondeleterious L1 insertion in an intron of the human factor VIII gene and further evidence of open reading frames in functional L1 elements. *Genomics.* 4:290–296.

10. Dombroski, B., S. L. Mathias, E. Nanthakumar, A. F. Scott, and H. H. Kazazian. 1991. Isolation of an active human transposable element. *Science (Wash. DC).* 254:1805–1808.

11. Mathias, S. L., A. F. Scott, H. H. Kazazian, J. D. Boeke, and A. Gabriel. 1991. Reverse transcriptase encoded by a human transposable element. *Science (Wash. DC).* 254:1808–1810.

12. Chamberlain, J. S., R. A. Gibbs, J. E. Ranier, P. N. Nguyen, and C. T. Caskey. 1988. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res.* 16:11141–11156.

13. Kitoh, Y., M. Matsuo, H. Nishio, T. Takumi, T. Nakajima, T. Masumura, J. Koga, and H. Nakamura. 1992. Amplification of ten deletion-rich exons of the dystrophin gene by polymerase chain reaction shows deletions in 36 of 90 Japanese patients with Duchenne muscular dystrophy. *Am. J. Med. Genet.* 42:453–457.

14. Matsuo, M., T. Masumura, T. Nakajima, Y. Kitoh, T. Takumi, H. Nishio, J. Koga, and H. Nakamura. 1990. A very small frame-shifting deletion within exon 19 of the Duchenne muscular dystrophy gene. *Biochem. Biophys. Res. Cummun.* 170:963–967.

15. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156–159.

16. Nakajima, T., M. Matsuo, Y. Kitoh, T. Takumi, H. Nishio, T. Masumura, J. Koga, and H. Nakamura. 1991. Screening of gene deletions by polymerase chain reaction in Japanese patients with Duchenne muscular dystrophy. *J. Neurol.* 238:6–8.

17. Scott, A. F., B. J. Schmeckpeper, M. Abdelrazik, C. T. Comey, B. O'Hara, J. P. Rossiter, T. Cooley, P. Heath, K. D. Smith, and L. Margolet. 1987. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. *Genomics.* 1:113–125.

18. Skowronski, J., and M. Singer. 1986. The abundant LINE-1 family of repeated DNA sequences in mammals: genes and pseudogenes. *Cold Spring Harbor Symp. Quant. Biol.* 51:457–464.

19. Roberts, R. G., D. R. Bentley, T. M. F. Barby, E. Manners, and M. Bobrow. 1990. Direct diagnosis of carriers of Duchenne and Becker muscular dystrophy by amplification of lymphocyte RNA. *Lancet.* 336:1523–1526.

20. Weiner, A. M., P. L. Deininger, and A. Efstratiadis. 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* 55:631–661.

21. Skowronski, J., T. G. Fanning, and M. F. Singer. 1988. Unit-length line-1 transcripts in human teratocarcinoma cells. *Mol. Cell. Biol.* 8:1385–1397.

22. Bratthauer, G., and T. G. Fanning. 1991. LINE 1 retrotransposon expression in human cancers. *Am. J. Hum. Genet.* 49(Suppl.):454.

23. Miki, Y., I. Nishio, A. Horii, Y. Miyoshi, J. Utsunomiya, K. Kinzler, K. Vogelstein, and Y. Nakamura. 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* 52:643–645.

24. Sakuraba, H., C. M. Eng, R. J. Desnick, and D. F. Bishop. 1992. Invariant exon skipping in the human α-galactosidase A pre-mRNA: a g + 1 to t substitution in a 5′-splice site causing Fabry disease. *Genomics.* 12:643–650.

25. Green, M. 1986. Pre-mRNA splicing. *Annu. Rev. Genet.* 20:671–708.

26. Matsuo, M., H. Nishio, Y. Kitoh, U. Francke, and H. Nakamura. 1992. Partial deletion of a dystrophin gene leads to exon skipping and to loss of an intra-exon hairpin structure from the predicted mRNA precursor. *Biochem. Biophys. Res. Commun.* 182:495–500.

27. Robberson, B., G. J. Cote, and S. Berget. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* 10:84–94.

28. Kontusaari, S., G. Tromp, H. Kuivaniemi, R. Ladda, and D. Prockop. 1990. Inheritance of an RNA splicing mutation (G + 1 IVS20) in the type III procollagen gene (COL3AI) in a family having aortic aneurysms and easy bruisability: phenotypic overlap between familial arterial aneurysms and Ehlers-Danlos syndrome type IV. *Am. J. Hum. Genet.* 47:112–120.

29. Nakajima, H., N. Kono, T. Yamasaki, K. Hotta, M. Kawachi, M. Kuwajima, T. Noguchi, T. Tanaka, and S. Tarui. 1990. Genetic defect in muscle phosphofructokinase deficiency. Abnormal splicing of the muscle phosphofructokinase gene due to a point mutation at the 5′-splice site. *J. Biol. Chem.* 265:9392–9395.

30. Korenberg, J., and M. Rykowski. 1988. Human genome organization: Alu, Lines, and the molecular structure of metaphase chromosome bands. *Cell.* 53:391–400.

31. Monaco, A. P., C. J. Bertelson, C. Colletti-Feener, and L. M. Kunkel. 1987. Localization and cloning of Xp21 deletion breakpoints involved in muscular dystrophy. *Hum. Genet.* 75:221–227.

32. Mager, D. L., P. S. Henthorn, and O. Smithies. 1985. A chinese Gδ + thalassemia deletion: comparison to other deletions in the human β-globin gene cluster and sequence analysis of the breakpoints. *Nucleic Acids Res.* 13:6550–6575.

33. Fitch, D. H. A., W. J. Bailey, D. A. Tagle, M. Goodman, L. Sieu, and J. L. Slightom. 1991. Duplication of the γ-globin gene mediated by L1 long interspersed repetitive elements in an early ancestor of simian primates. *Proc. Natl. Acad. Sci. USA.* 88:7396–7400.

34. Love, D. R., S. B. England, A. Speer, R. F. Marsden, J. F. Bloomfield, A. L. Roche, G. S. Gross, R. C. Mountford, T. J. Smith, and K. E. Davies. 1991. Sequences of junction fragments in the deletion-prone region of the dystrophin gene. *Genomics.* 10:57–67.

35. Lehman, M. A., W. J. Schneider, T. C. Sudhof, M. S. Brown, J. L. Goldstein, and D. W. Russell. 1985. Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science (Wash. DC).* 227:140–146.

36. Koenig, M., A. P. Monaco, and L. M. Kunkel. 1988. The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell.* 53:219–228.