Am. J. Hum. Genet. 63:436–446, 1998

A Novel *Alu*-Like Element Rearranged in the Dystrophin Gene Causes a Splicing Mutation in a Family with X-Linked Dilated Cardiomyopathy

Alessandra Ferlini,[1] Nazzareno Galié,[3] Luciano Merlini,[4] Caroline Sewry,[1,2] Angelo Branzi,[3] and Francesco Muntoni[1]

[1]Department of Paediatrics & Neonatal Medicine, Imperial College School of Medicine, and [2]Muscle Cell Biology Group, MRC Clinical Sciences Centre, Hammersmith Hospital, London; and [3]Istituto di Malattie dell'Apparato Cardiovascolare, Università di Bologna, and [4]Istituto Ortopedico Rizzoli, Bologna

Summary

We have identified and characterized a genomic sequence with some features typical of *Alu*-like mobile elements rearranged into the dystrophin gene in a family affected by X-linked dilated cardiomyopathy. The *Alu*-like sequence rearrangement occurred 2.4 kb downstream from the 5′ end of intron 11 of the dystrophin gene. This rearrangement activated one cryptic splice site in intron 11 and produced an alternative transcript containing the *Alu*-like sequence and part of the adjacent intron 11, spliced between exons 11 and 12. Translation of this alternative transcript is truncated because of the numerous stop codons present in every frame of the *Alu*-like sequence. Only the mutant mRNA was detected in the heart muscle, but in the skeletal muscle it coexisted with the normal one. This result is supported by the immunocytochemical findings, which failed to detect dystrophin in the patient's cardiac muscle but showed expression of a reduced level of protein in the skeletal muscle. Comparative analysis of the *Alu*-like sequence showed high homology with other repeated-element–containing regions and with several expressed sequence tags. We suggest that this *Alu*-like sequence could represent a novel class of repetitive elements, reiterated and clustered with some known mobile elements and capable of transposition. Our report underlines the complexity of the pathogenic mechanism leading to X-linked dilated cardiomyopathy but suggests that differences in tissue-specific expression of dystrophin mutations may be a common feature in this condition.

Received October 21, 1997; accepted for publication May 15, 1998; electronically published June 29, 1998.

Address for correspondence and reprints: Dr. Alessandra Ferlini, Neuromuscular Unit, Department of Paediatrics & Neonatal Medicine, Hammersmith Hospital, Du Cane Road, London W12 ONN, United Kingdom. E-mail: aferlini@rpms.ac.uk

0002-9297/98/6302-0019$02.00

Introduction

The cardiomyopathies represent a major cause of cardiovascular morbidity and mortality, and, within this group, dilated cardiomyopathy (DC) is the most common of the various forms (Manolio et al. 1992). DC can result from numerous causes, but increasing interest in the genetic predisposition to this condition recently has developed. Although several disease loci responsible for the dominant forms of DC have been mapped (Durand et al. 1995; Krajinovic et al. 1995; Olson and Keating 1996; van der Kooi et al. 1997), the only gene known to be responsible for one form of DC is the dystrophin gene located on the X chromosome (MIM 310200). Mutations in this gene are usually associated either with the severe Duchenne-type muscular dystrophy (DMD) or with its milder allelic variant, Becker muscular dystrophy (BMD) (for review, see Ahn and Kunkel 1993), in which the heart also is affected. In rare instances, however, dystrophin mutations have been associated with X-linked DC (XLDC) in the absence of muscle weakness (Muntoni et al. 1993; Towbin et al. 1993; Yoshida et al. 1993; Milasin et al. 1996). The only suggestion of skeletal-muscle involvement in these families was the presence of either raised serum creatine kinase (CK) levels (Muntoni et al. 1993; Towbin et al. 1993; Yoshida et al. 1993) or occasional myalgias after strenuous physical exercise (Milasin et al. 1996).

Interspersed repetitive DNA elements (IRS) represent different classes of DNA sequences, with high, medium, and low grades of reiteration in the eukaryote genome (Weiner et al. 1986; Deininger et al. 1992; McNaughton et al. 1995). Three major types of IRS have been identified—the long interspersed nucleotide elements (LINEs; ~6 kb long), the short interspersed nucleotide elements (SINEs; ~300 bp), and the transposon-like human elements (THE; 2.5–7 kb). These three classes of IRS, together with many other, less abundant, repetitive elements, have provided an important contribution to genome evolution (Makalowski et al. 1994). These elements, also called "mobile elements," possess the capability to change their genomic location by transposi-

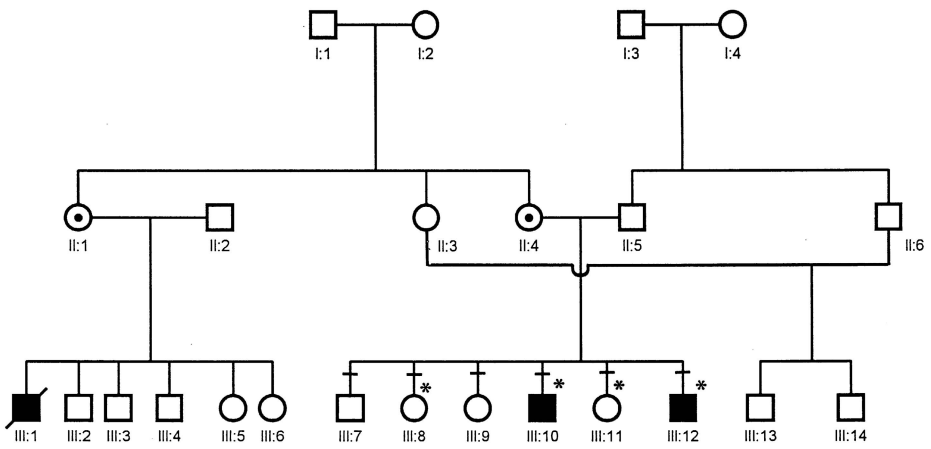

Figure 1 Pedigree of family with XLDC. The blackened squares denote affected patients; and the circles including black dots denote obligate female carriers. A horizontal tic mark above a symbol indicates that the individual was analyzed; and an asterisk indicates that the individual is a carrier of the dystrophin rearrangement.

tion. This represents an incompletely understood event by which a master repetitive element, representing the ancestor consensus sequence from which the different IRS families have originated, moves its genomic localization via an RNA intermediate and reverse transcription (Mathias et al. 1991; Craig 1995). Among the SINEs, the *Alu* family constitutes the major interspersed repeat of primate genomes, with >500,000 copies in the human haploid genome (Britten et al. 1988).

Alu elements and LINEs occasionally have been associated with genetic diseases, mainly by providing the sites for homologous recombination, because of their high homology to master elements. These recombinations result in deletions or other, more-complex genomic rearrangements (Muratani et al. 1991; Wallace et al. 1991; Li and Bray 1993; Vidaud et al. 1993; Janicic et al. 1995). Insertions of IRS into human genes via retroposition also have been hypothesized but, both in the human and in other species, have been demonstrated only in a few cases (Homes et al. 1994; Kingsmore et al. 1994). Another rare mutation mechanism can be a point mutation that activates both a cryptic splice site and the splicing of intronic *Alu* elements into the mature transcript (Mitchell et al. 1991; Knebelmann et al. 1994).

In this report we present evidence for a mutation in the dystrophin gene in a family with XLDC due to a rearrangement of an *Alu*-like element within intron 11. This event modified the splicing in a tissue-specific manner, resulting in absence of dystrophin expression only in the heart muscle. Comparative sequence analysis suggests that this *Alu*-like sequence is part of a novel, transcriptionally active repetitive-element family, given the absence of significant homology with *Alu,* LINEs, and other known IRS consensus sequences.

Subjects and Methods

Clinical Studies

This previously unreported Italian family is of Sardinian descent. Diagnosis of DC was established on the basis of clinical history, physical examination, chest radiography, electrocardiography (EKG), echocardiography, and cardiac catheterization. The diagnostic criteria for DC were those promulgated by the World Health Organization (Richardson et al. 1996). The proband (individual III-12; fig. 1) is a 20-year-old male who was well and physically very active until the age of 16 years, when he was admitted after contracting viral hepatitis. Radiography of the chest showed an unexpected moderate cardiomegaly, and high serum CK levels were documented (1,970 IU/liter [normal <195 IU/liter]). EKG showed sinus tachycardia, left-atrial enlargement, and right-bundle-branch block. The left-ventricular ejection fraction (EF) was reduced (17%), as assessed by echocardiography. At that time the patient was virtually asymptomatic, and a diagnosis of DC was made. One year later the patient was admitted for an episode of acute congestive heart failure, and echocardiography confirmed both the severe dilatation of the right and left ventricles and a reduction of the EF. A right-ventricular endomyocardial biopsy and a skeletal-muscle biopsy were performed, for diagnostic purposes, after informed consent was obtained. Since then, the patient has been in a stable clinical condition and complains only of dyspnea, after strenuous exercise. The results of a neurological examination in 1996 were unremarkable; in particular, the patient had normal muscle strength and did not have muscle wasting or hypertrophy. Several other family members were affected by DC, and analysis of the

Table 1

Primer Sequences Used

Primer	Sequence (5′–3′)	Orientation	Annealing Temperature (°C)
Ex11F	GAAGTAAGCTGATTGGAAC	Forward	61
Ex12R	GTTGTTGTACTTGGCGTTTTA	Reverse	61
Int11F	GATGTTCAGTAATAAGTTGCTTT	Forward	60
Int11R	AGCTTCCAAAACTTGTTAGTC	Reverse	60
SeqF	ACAGGGTTTGGATAGATCCAGTC	Forward	60
SeqR	GACTGGATCTATCCAAACCCTGT	Reverse	60
Dysint	GTTACTTCAGATGATGACTCATTG	Reverse	61
Trasp1A	CAAACCTTTTCTACAAGAAATGGT	Forward	62
Trasp1B	ACCATTTCTTGTAGAAAAGGTTTG	Reverse	62
Dysgen	GTTCAGTTTCCATTATCATGGTC	Forward	63

pedigree suggested an X-linked inheritance (fig. 1). Specifically, the 29-year-old older brother (individual III-10; fig. 1) also was found to have elevated CK levels. The results of physical examination of this individual were unremarkable, whereas EKG showed ventricular hypertrophy. Echocardiography disclosed a left-ventricular dilatation, reduced (40%) left-ventricular EF, and mild mitral-valve incompetence. The individual is virtually asymptomatic. A maternal cousin (individual III-1) died at the age of 18 years because of heart failure. In addition to the cardiac involvement, he also had muscle weakness suggesting a BMD phenotype. Individuals III-7, III-8, III-9, and III-11 are clinically normal and have normal serum CK levels.

Cardiac-muscle and skeletal-muscle biopsy specimens from the proband (individual III-12; fig. 1) were available for immunocytochemical and transcription studies. Peripheral blood was collected from six members of the family (individuals III-7–III-12; fig. 1).

Immunocytochemical Study

Skeletal-muscle and cardiac-muscle biopsy specimens were studied according to standard techniques, including immunocytochemistry (Sewry et al. 1993). Six-micron unfixed cryostat sections were immunostained with a panel of anti-dystrophin antibodies (Novocastra).

Molecular Analysis

Transcription studies.—Total RNA was isolated from skeletal-muscle and cardiac-muscle biopsy specimens by the method of Chomczynsky and Sacchi (1987). Reverse-transcription–PCR for cDNA synthesis was performed with random hexanucleotide primers and the Moloney murine leukemia virus reverse transcriptase (Pharmacia), according to a procedure that has been described elsewhere (Muntoni et al. 1995).

The amplification products were electrophoresed on a high-resolution agarose gel (Electran; Sigma) and were visualized after being stained with ethidium bromide. The amplification products were purified on Glass Max columns (Gibco), if necessary. The PCR reactions (Thermal Cycle model 480; Perkin-Elmer) were performed by use of Taq Pfu Polymerase (Stratagene); the amplification conditions are available on request.

Genomic studies.—Genomic DNA was isolated from peripheral blood obtained from individuals III-7–III-12 (fig. 1), by standard methods (Kunkel et al. 1983). Multiplex DNA amplifications to exclude the common deletions of dystrophin exons were performed (courtesy of Prof. M. A. Melis), according to the methods of Chamberlain et al. (1988) and Beggs et al. (1990).

The pairs of oligonucleotides Ex11F/Int11R and Int11F/Ex12R (table 1) were designed for amplification and sequencing of the exon-intron boundaries of intron 11, according to the published sequences (accession number I07692.gb_pat; Monaco et al. 1988). Several oligonucleotides also were designed to amplify the genomic region including the 5′ end of intron 11 in controls and patients. The sequences of all oligonucleotides used are shown in table 1; the conditions of amplification are available on request. For PCR, Perkin-Elmer DNA Thermo Cycler 480 and Perkin-Elmer 2400 were used. *Taq* Plus polymerase (Stratagene) and *Taq* Expand System (Boehringer) were used for the amplification reactions, with the appropriate incubation buffers. Amplification products were gel purified, when necessary, by agarase digestion (Boehringer).

To isolate the upstream region of intron 11 in individual III-12, we adopted a PCR-walking method enabling us to amplify and sequence the 5′ 2.5-kb intronic region (Ferlini and Muntoni 1998). Several attempts, by means of Long *Taq* polymerases approaches (*Taq* Expand System and Takara LA polymerase), to amplify the genomic fragment downstream from the *Alu*-like sequence rearrangement failed, probably because of the large size of intron 11.

After digestion of genomic DNA from patients and controls, Southern blot analysis was performed with *Bgl*II, *Hin*dIII, *Xmn*I, *Eco*RI, and *Taq*I restriction en-

zymes, and the products were transferred onto Gene Screen Plus (NEN) in 10 × SSC buffer. Membranes were hybridized by α-[^{32}P] (dATP)–radiolabeled probes (Megaprime and Nick Translation kits; Amersham) by the Church and Gilbert (1984) method, under highly competitive conditions (300 mg herring-sperm DNA). The probes used were the 2b-3 dystrophin cDNA fragment (a gift from Dr. R. Roberts), the 159-bp sequence inserted into individual III-12's RNA (probe Past), and a 2.5-kb intron 11 amplification product (probe 2.5 KM) obtained from individual III-12. For the Past and 2.5 KM probes, very-high-stringency conditions (hybridization at 68°C and washing at 70°C) were used. Membranes were exposed for 1–5 days at −80°C, for autoradiography.

Cloning and sequencing.—For cloning procedures, the cDNA-amplification products were cloned in the pGEM vector (TA cloning system; Promega) and were sequenced manually (Sequenase kit, version 2.0; Pharmacia). Genomic DNA fragments were cloned into PCR Script SK$^+$ AMP Vector (Stratagene), and this was followed by either manual (Sequenase kit, version 2.0; Pharmacia) or automated (Applied Biosystems) sequencing.

Comparative sequence analysis.—Comparative sequence analysis was done by means of the GCG Wisconsin Package, with FASTA and BLAST.

Results

Skeletal-Muscle and Cardiac-Muscle Biopsies

Histological analysis of the skeletal-muscle biopsy specimen from individual III-12 (see the pedigree in fig. 1) showed a mild myopathic picture, with mild variability in fiber size, an increased number of internal nuclei, a few areas of degeneration/regeneration, and no fibrosis (fig. 2*e*). In contrast, a severe dystrophic cardiomyopathy, characterized by marked fibrosis and variability in fiber size, was present in the cardiac-muscle biopsy specimen (fig. 2*f*).

Immunocytochemical analysis of individual III-12's skeletal-muscle biopsy specimen showed a moderate reduction in staining when a panel of anti-dystrophin antibodies was used (fig. 2*b*). The cardiac muscle, on the contrary, showed a total absence of staining (fig. 2*d*). These results indicated the presence of a dystrophinopathy that affected the cardiac muscle more severely than it affected the skeletal muscle.

Transcription Studies

Transcription analysis of total skeletal-muscle RNA of the proband, III-12, performed with Ex11F and Ex12R used as primers to amplify exons 11 and 12, showed a higher-molecular-weight fragment, in addition to that having the expected size (fig. 3*a*). In contrast, this larger product was virtually the only one present in the cardiac-muscle biopsy specimen from the propositus, and it was not observed in five skeletal-muscle and three cardiac-muscle tissues from controls (fig. 3*a*). Sequence analysis of these two fragments demonstrated that the smaller-size product represented the normal gene transcript, whereas the other fragment contained an inserted sequence of 159 bp. A schematic representation of the alternative dystrophin transcript is shown in figure 3*b*.

The 159-bp sequence inserted into the patient's mRNA caused a nontranslated message, since analysis of its sequence showed 12 stop codons (one every 13 bp), which were present in all three reading frames and were responsible for a premature stop of the message, despite the conservation of the canonical frame of the dystrophin message. Furthermore, the sequence contained 39% G+C and 61% A+T, indicating that this is an AT-rich region. These characteristics, together with genomic data and comparative analysis (see below), are suggestive of an *Alu*-like element spliced into the transcript and causing a prematurely truncated dystrophin. The finding that the cardiac muscle expressed only the truncated transcript whereas, in the skeletal-muscle RNA, the latter was coexpressed with the normal transcript is in keeping with both the immunocytochemical and clinical findings in this family, in which no significant muscle weakness was present.

Genomic Studies

PCR and sequence analysis.—No genomic deletion of the dystrophin gene was found after combined multiplex PCR analysis (data not shown). Sequence analysis of the canonical donor and acceptor splicing sites of intron 11 in the propositus, individual III-12, excluded a mutation in both these regions, including the flanking splicing consensus sequences (data not shown).

A pair of primers located within the transcribed *Alu*-like sequence—namely, SeqF and Dysint (table 1 and fig. 3*c*)—were designed and used to amplify both skeletal-muscle and cardiac-muscle cDNAs and genomic DNA from patients and relatives (fig. 4*a*). This confirmed that the same *Alu*-like fragment spliced into the patient's RNA was present in the genomic DNA of the two affected brothers (III-10 and III-12) and in that of two of the three sisters (III-8 and III-11) but not in that of the unaffected brother (III-7). Similarly, when both the same reverse primer (Dysint) and a forward primer (Dysgen) located in the intronic region flanking, upstream, the *Alu*-like sequence (fig. 3*d*) were used, an amplified fragment present only in the patients and not in 30 female controls was found (fig. 4*b*). Further amplification, using the same forward primer and a different reverse primer (Trasp1B), located in the middle of the *Alu*-like sequence (fig. 3*c*), gave identical results (fig. 4*c*), indicating the presence of a rearrangement in this intronic region.

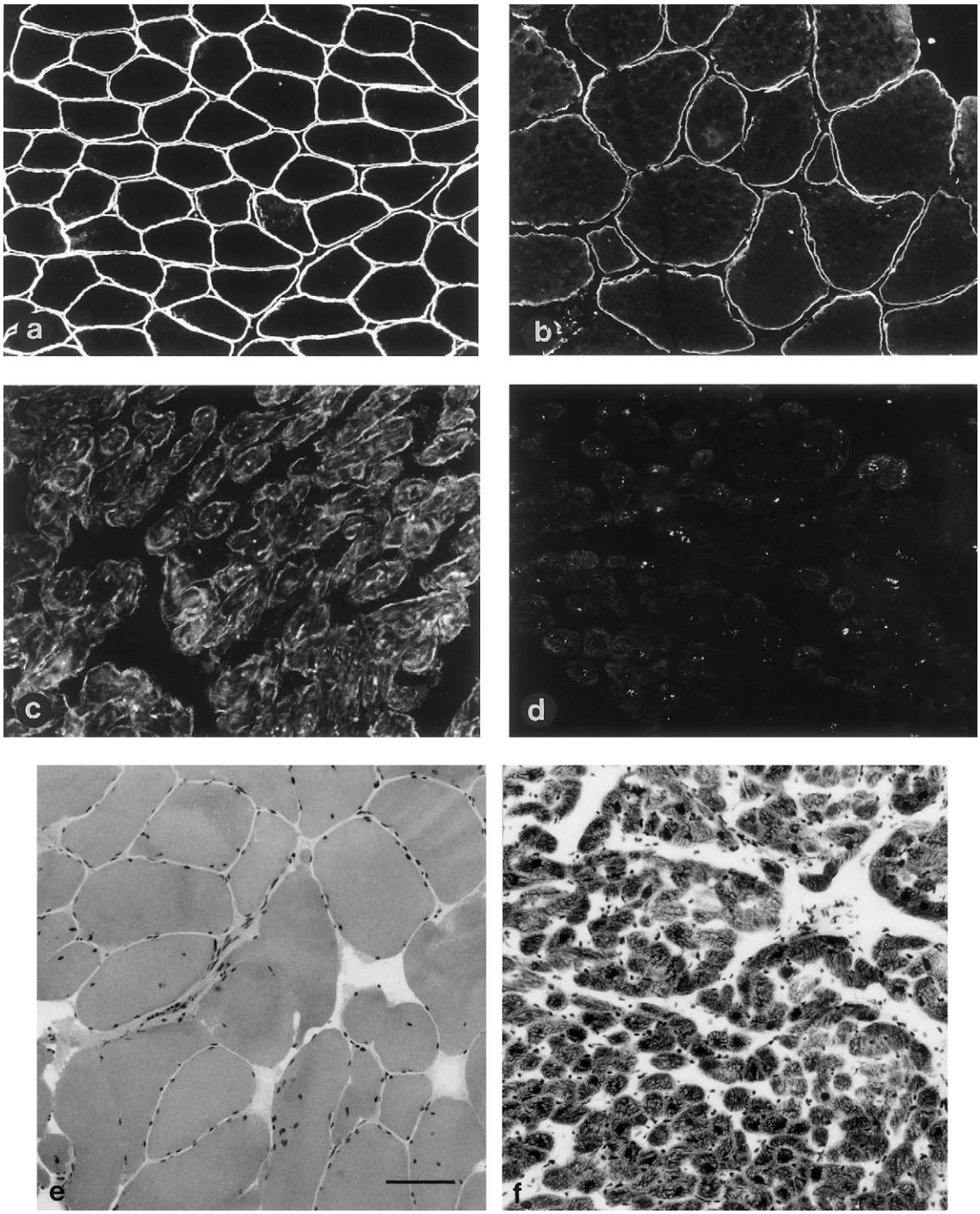

Figure 2 *a–d,* Immunocytochemical analysis of skeletal-muscle and cardiac-muscle biopsy specimens from a control and individual III-12, by means of Dys3, an N-terminal anti-dystrophin antibody. Expression of dystrophin is normal in skeletal muscle and cardiac muscle of the control (*a* and *c*). In individual III-12's skeletal-muscle DNA (*b*), there is a reduction in staining, compared with that in the control, whereas, in his cardiac-muscle (*d*), dystrophin is absent. *e* and *f,* Section of skeletal-muscle biopsy specimen (*e*) and cardiac-muscle biopsy specimen (*f*), stained with hematoxylin and eosin (scale bar = 100 μm). The specimens show myopathic changes in the skeletal muscle and fibrosis and cardiomyocyte hypotrophy in the cardiac muscle.

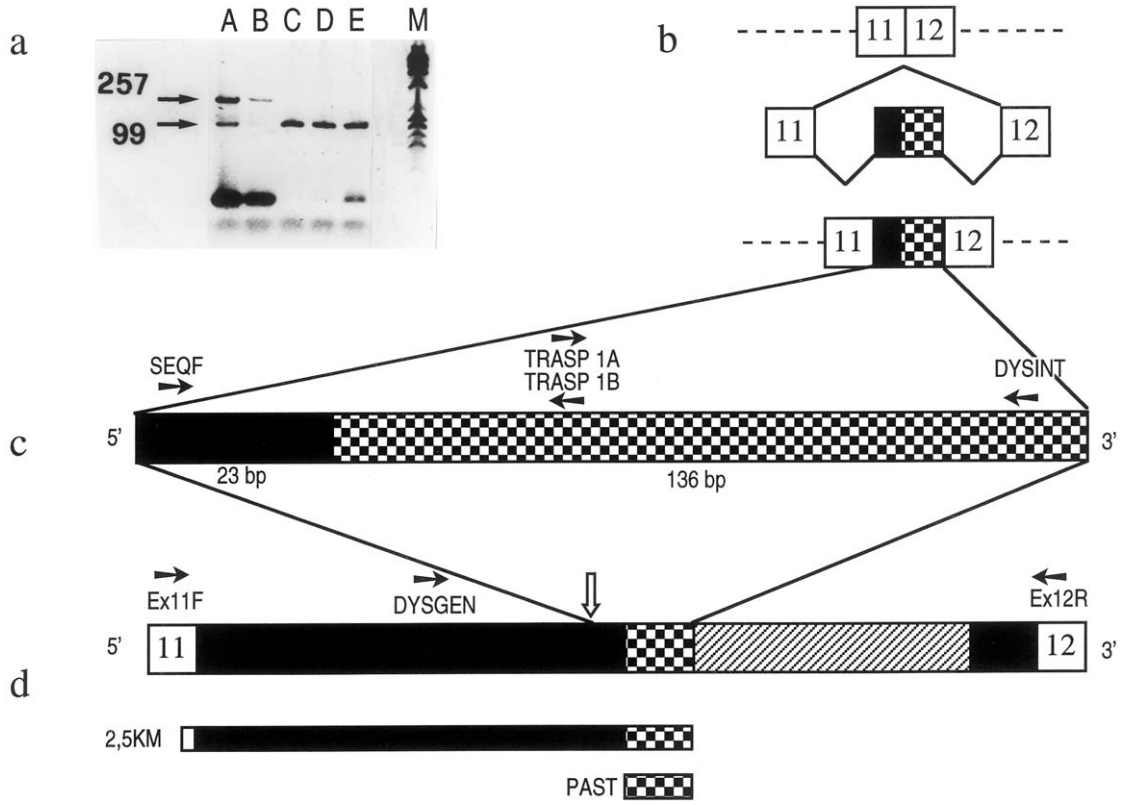

Figure 3 *a,* Transcription analysis of skeletal-muscle RNA (lane A) and cardiac-muscle RNA (lane B) of individual III-12, performed with Ex11F and Ex12R used as oligonucleotides. A 257-bp, larger fragment is visible, in addition to the 99-bp fragment expected. The alternative transcript represented the only transcript in the cardiac muscle of the patient and is not present in skeletal-muscle RNA (lanes C and D) or cardiac-muscle RNA (lane E) of a control. Lane M, Marker IX (Boehringer). The smaller bands (<50 bp) represent primer dimers. *b,* Schematic representation of both the alternative and the normal dystrophin transcripts. The unblackened squares denote the normal exons 11 and 12; and the blackened squares and the patterned squares denote the intronic region and the *Alu*-like sequence, respectively. *c,* Schematic representation of the mRNA configuration in individual III-12. The rightward- and leftward-facing arrows indicate, respectively, the forward and reverse primers used for cDNA and genomic-DNA amplification; the blackened bar denotes the normal intron 11 region (23 bp), spliced in the transcript; and the patterned bar denotes the *Alu*-like sequence (136 bp). *d,* Schematic representation of the dystrophin gene intron 11 in individual III-12. The rightward- and leftward-facing arrows indicate, respectively, the forward and reverse primers used for genomic-DNA amplification; the unblackened squares indicate exon 11 and 12 of the dystrophin gene; the blackened bar denotes the normal intron 11 region; the patterned bar denotes the *Alu*-like sequence; and the diagonally striped bar denotes the intron 11 region whose sequence is still unknown; and the unblackened downward-facing arrow indicates the cryptic splicing site normally present in intron 11. The two genomic probes (2.5 KM and Past) used for Southern blot analysis are shown below the diagram.

However, amplification under very-low-stringency conditions (58°C) of the 3′ region of the *Alu*-like sequence (fig. 3*c*), with the forward primer Trasp1A and the reverse primer Dysint, resulted in an amplification in all family members and in controls, including a child with both DMD and a genomic deletion encompassing intron 11 (data not shown). This finding suggests that this sequence is present in at least one copy in the human genome.

All these data suggest that a rearrangement of a novel *Alu*-like sequence in intron 11 of the dystrophin gene occurred in the family that we studied, segregating with the disease. We can hypothesize that this rearrangement also was present in individual III-1, who also suffered with muscle weakness and DC; and, if this is the case, we can speculate that this rearrangement in intron 11 could have occurred in the maternal grandmother (I-2; fig. 1), representing an obligate carrier. A schematic representation of the genomic structure of this region in our patients, as well as the location of the primers and the probes used, is shown in figure 3*d*.

To determine the site of the *Alu*-like rearrangement in intron 11, we used a forward primer located in exon 11 (Ex11F) and a reverse primer located in the 3′ end of the sequence (Dysint). We amplified a 2,552-bp fragment from patients' genomic DNA but not from that of six male controls (fig. 4*d*). To refine the position in which the *Alu*-like insertion occurred in intron 11, we amplified and sequenced this fragment by PCR walking, using various sets of primers (Ferlini and Muntoni 1998), and

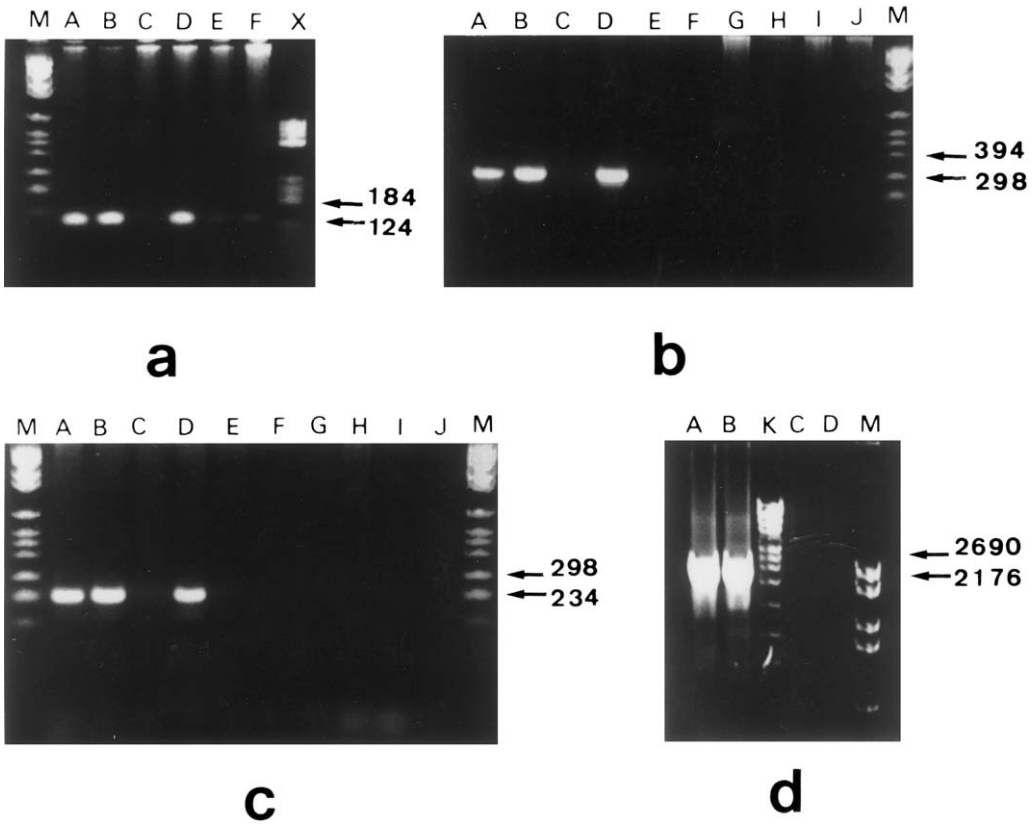

Figure 4 Amplification of the intron 11 region containing the *Alu*-like rearrangement. *a,* Amplification, by primers SeqF and Dysint, of genomic DNA from individual III-10 (lane A), individual III-12 (lane B), female III-9 (lane C), and female III-11 (lane D), and skeletal-muscle and cardiac-muscle cDNAs from individual III-12 (lanes E and F, respectively). A 137-bp fragment was amplified from all samples, with the exception of that from individual III-9 (a noncarrier sister of the propositus). Lane X, Molecular-weight marker V (Boehringer). *b,* Amplification, by primers Dysgen (located in the intronic region flanking upstream the *Alu*-like sequence) and Dysint, of genomic DNA from affected males III-10 and III-12 (lanes A and B, respectively), one noncarrier and one carrier sister (individuals III-9 [lane C] and III-11 [lane D], respectively), one healthy brother (individual III-7 [lane E]), and healthy controls (lanes F–J). The 301-bp fragment is present only in the patients and in a carrier sister and not in the healthy brother, the noncarrier sister, or controls. Lane M, Molecular-weight marker VI (Boehringer). *c,* Amplification of genomic DNA from the same individuals as are represented in *b,* by the same forward primer (Dysgen) but with a different reverse primer (Trasp1B). The pattern of amplification (236 bp) is identical to that obtained with the primer combination described in *b,* excluding the possibility that the amplification obtained in this family is due to a point mutation occurring in the region covered by the Dysint primer. Lane M, Molecular-weight marker VI (Boehringer). *d,* Amplification, by primers Ex11F and Dysint, of genomic DNA from the two affected brothers, III-10 and III-12 (lanes A and B, respectively), and from two controls (lanes C and D). A fragment that is ~2.5 kb is present only in the patients. Lanes K and M, Molecular-weight markers X and VI, respectively (Boehringer).

identified a minimal region of 22 bp, at nucleotide positions 2349–2371 of intron 11, normally present in patients and controls (data not shown).

In conclusion, we were able to establish that the 2.5-kb fragment contains both the 5′ 2,371-bp of intron 11 and a 136-bp *Alu*-like sequence in which a rearrangement occurred. Interestingly, the 3′ end of the novel intron 11 sequence, which is flanking, upstream, the *Alu*-like sequence, contains a conserved AG acceptor splicing site with a flanking pyrimidine-rich consensus sequence, preferentially utilized in the patient's cardiac muscle.

Southern blot analysis.—Hybridization of *Bgl*II and *Hin*dIII Southern blots, using the cDNA 2b-3 as a probe encompassing exon 11 and 12 of the dystrophin gene, failed to show any gross rearrangements in patients' DNA. On the other hand, sequence analysis of the 5′ end of intron 11 showed both *Hin*dIII and *Bgl*II restriction sites in the region upstream from the rearrangement, providing an explanation for the normal pattern found in patients. Hybridization of *Bgl*II-, *Hin*dIII-, *Xmn*I-, *Taq*I-, and *Eco*RI-digested Southern blots, using the probe Past (the 159-bp PCR fragment; fig. 3*d*), gave hybridization patterns similar to that obtained with a multicopy-probe hybridization (data not shown). Hybridization with the *Xmn*I digestion's membrane, by use of the 2.5-kb PCR fragment (probe 2.5 KM) including both the 5′ end of intron 11 and the *Alu*-like sequence (probe Past) (fig. 3*d*) under high-stringency conditions

(68°C–72°C), showed the presence of a unique fragment (~10 kb) in the two affected brothers as well as in one sister, who also carried the normal-size allele (fig. 5). The pattern of segregation of the *Xmn*I fragment was identical to that found in the genomic amplifications mentioned above, indicating that the two brothers and two sisters (individuals III-10 and III-12 and individuals III-8 and III-11, respectively; see fig. 1) are carrying the same genomic rearrangement, whereas the healthy brother (III-7) and the other sister (III-9) are not. Because of the large size (>30 kb, as reported in the Leiden Muscular Dystrophy Pages) of the 3′ of intron 11, we have not been able to ascertain whether the donor splice site has been created by the genomic rearrangement or is endogenously present in intron 11; nor have we been able to ascertain the 3′ end of the rearrangement. We cannot exclude the possibility that the *Alu*-like sequence has been inserted by retrotransposition in intron 11. We have, however, established that a normally present acceptor splicing site in intron 11 of the dystrophin gene was activated by a rearrangement consisting of $\geqslant$100 bp and involving an *Alu*-like sequence.

Comparative sequence analysis.—We have submitted both the wild-type 5′ end of intron 11, nucleotide positions 1–2371 (accession number Y13187), and the rearranged 136-bp *Alu*-like sequence (accession number Y13186) to the DDBJ/EMBL/GenBank nucleotide-sequence databases.

Comparative analysis of the *Alu*-like inserted sequence failed to detect any significant homology either with the *Alu* and LINE1 consensus, with other known repetitive-element families, or with THE1; instead, the analysis showed a significant homology (>80%, on a minimum overlap of 100 bp) with several human genomic sequences. These included (*a*) a series of cosmids mapped on X chromosome—in Xq28 (accession number Z47066), in Xq21-q22 (accession numbers Z68332, Z70050, Z70758, Z73967, Z80900, Z69367, Z75894, Z68871, Z68328, Z70689, Z69838, Z74477, Z68871, Z70226, Z68289, Z73497, Z75745, and Z70233), in Xq23 (accession numbers AC000055 and AC000114), in Xq22 (in the expressed sequence tags [ESTs] of glycerol kinase pseudogene [accession number Z73986]), and in Xp22 (accession number U70984); (*b*) a region flanking the 3′ UTR of the FRAXA gene on Xq28 and containing several *Alu* elements and LINE1s; (*c*) cosmids Z49862, Z49235, and Z68165, flanking the Huntington gene on chromosome 4p16.3; and (*d*) other cosmids, located in 7q21-q22 (accession numbers AC000066, AC000121, AC000119, and AC000057) and including the BRCA2 gene region on 13q12-q13 (accession number Z73360). The homologous region that our *Alu*-like sequence shared with these sequences always occurred within nucleotide positions 25–136 or 25–79. Interestingly, this was a region in which a significant (80%) homology with two eukaryotic promoters was found (accession numbers EPD14029 and EPD17097). In addition, comparative analysis with the EMBL-EST bank revealed homology with 13 ESTs, the most relevant (80%) being two *Homo sapiens* cDNA sequences (accession numbers C02353 and Z25899). All these ESTs represent truncated messages and do not have an open reading frame.

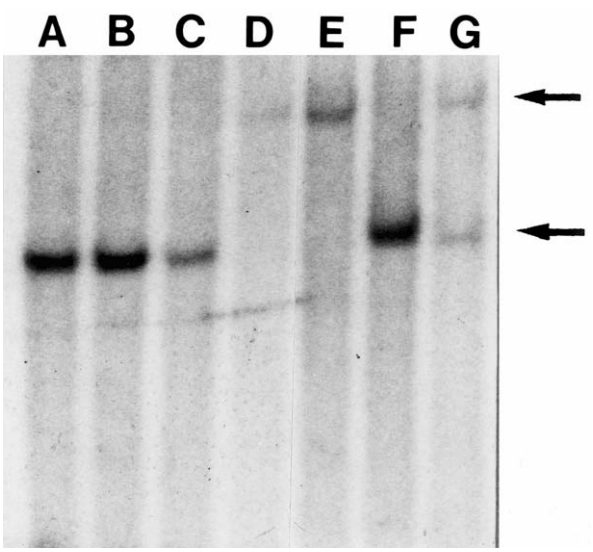

Figure 5 Southern-blot hybridization using both the 2.5 KM fragment as a probe and *Xmn*I restriction enzyme. Lanes A–C, Normal DNA from controls. Lanes D and E, DNA of the affected brothers, III-10 and III-12. Lane F, DNA from female III-9. Lane G, DNA from female III-11. In addition to the normal fragment, an abnormal fragment that is ~10 kb is present only in the two affected males and the carrier sister, III-11.

Comparative analysis of the 5′ region of intron 11 of the dystrophin gene indicated a significant (70%) homology with several cosmids mapped at Xq22 and 7q22, as well as with *Homo sapiens* ESTs (Ferlini and Muntoni 1998). In this case, again, the cosmids contained several *Alu* elements and LINE1s.

Discussion

In this report we have described a family affected by XLDC with an unusual dystrophin gene mutation. The disease mechanism in this family was an intronic rearrangement inducing the inclusion of an *Alu*-like sequence and part of the 5′ adjacent intronic region in the dystrophin mRNA. This sequence contained several stop codons, causing a prematurely truncated message. In skeletal muscle, the normal transcript coexisted with the alternative message, allowing the production of sufficient dystrophin to account for the lack of significant muscle damage and weakness. On the other hand, the

alternative transcript represented virtually the only message in the cardiac muscle and is likely to be responsible for the isolated cardiomyopathic phenotype.

The genomic mutation in this family consisted of a small rearrangement located 2.4 kb downstream from the 5′ end of intron 11 of the dystrophin gene. Both transcription and genomic data suggest that the *Alu*-like–sequence insertion is responsible for the XLDC in our patients. The fact that the mRNA insertion introduced different stop codons in the canonical reading frame demonstrates that the spliced sequence does not represent a common alternative spliced form of the dystrophin gene, as is also evident from the analysis of several samples of skeletal-muscle and cardiac-muscle RNAs from controls. Moreover, the genomic segregation of the rearranged sequence with the XLDC in the members of this family whom we analyzed confirms the linkage between this mutation and the disease. Amplification of DNA from several controls (>80 chromosomes) failed to show this genomic configuration, confirming that it does not represent a common DNA polymorphism.

This rearrangement contained a sequence with similarities both to SINE sequences, including *Alu,* and to LINE1 (AT content, multiple stop codons' contents, homology with EST sequences, and reiterated presence, in both strand orientations, in the human genome). Furthermore, low-stringency-condition amplification, as well as Southern analysis, gave a multicopy sequence pattern.

The comparative data analysis suggested that this *Alu*-like sequence represents a novel expressed multicopy genomic element. This hypothesis is supported by the fact that this sequence is (*a*) present in multiple copies in the human genome; (*b*) homologous to other sequences in the data bank, in a way that is suggestive of a family of interspersed repeats; and (*c*) homologous to several human ESTs, all of them without an open reading frame. This strongly suggests that these ESTs represent RNA intermediates, originating from related genomic sequences, transcriptionally active, and potentially capable of retrotransposition. In addition, the genomic regions with homology to our sequence are rich in *Alu* elements, THEs, LINE1s, and other repetitive elements, which normally are clustered as a well-recognized feature of IRS (Britten 1996; Weiner et al. 1986).

Both the evidence of a rearrangement in this family and the features of this novel repetitive sequence suggest that the latter might be part of a mobile element. This element could have been rearranged into intron 11 because of a nonhomologous recombination, facilitated by the presence of the repeats in this region, or it could have been retrotransposed and inserted via an RNA intermediate. This latter possibility also seems to be supported by the fact that the genomic as well as the EST regions, described above and sharing homology with the *Alu*-like sequence, have both strand orientations. Additional evidence supporting the hypothesis that this sequence meets the basic requirements of IRS (Britten 1996) derives from the observation that homology with other known repetitive elements is present in intron 11 of the dystrophin gene.

Alu or LINE1 insertion rearrangements have been reported to be associated with human diseases only in very rare instances (Muratani et al. 1991; Wallace et al. 1991; Li and Bray 1993; Vidaud et al. 1993; Janicic et al. 1995). *Alu*, THE, and LINE1 also have been identified in dystrophin introns and, rarely, are associated with gross mutations leading to DMD or BMD (Pizzuti et al. 1992; Narita et al. 1993; Homes et al. 1994). In an earlier report, a splicing mutation due to an insertion of the 3′ region of a LINE1 in exon 44 of the dystrophin gene was described in a DMD patient. This insertion determined a larger exon and also its skipping, without affecting the canonical splicing site. However, the insertion mutation created a sequence with only one nucleotide (position +4) different from the flanking consensus sequence (Narita et al. 1993). As reported for other human and nonhuman genes, mutations either (*a*) abolishing or reducing the efficiency of a canonical splicing site or (*b*) creating a cryptic one may result in preferential splicing regulation of alternative transcripts (Nelson and Green 1990; Hahn et al. 1995; Sebillon et al. 1995).

In the family that we have described, the insertion of this novel *Alu*-like element activates a cryptic acceptor splicing site, already present in intron 11 of the dystrophin gene, without any mutation in either the canonical splicing sites or the consensus sequences. We have demonstrated that this mutation affects the splicing in a tissue-specific way. To the best of our knowledge, this is the first report in which such a mechanism has been reported in a human genetic disease. However, the possibility that LINE, *Alu,* and other IRS might affect and control the transcription of adjacent genes in various species as well as in the human has been suggested in a previous report (Favor and Morawetz 1992). Furthermore, in a few instances, although never in the human, a tissue-specific induction of gene expression has been found as a consequence of an IRS sequence's insertion (Adeniyi-Jones and Zasloff 1985; Tanda and Corces 1991). It seems most likely that this *Alu*-like sequence might have played a direct role in regulation of the splicing of the dystrophin gene in this family with XLDC. A possible explanation is that the rearrangement that occurred in intron 11 has brought the novel *Alu*-like sequence (and probably other flanking regulating regions) close to the cryptic splice site, which is normally present in the intron. This event activated this latter site, leading to the alternative, nonefficient splicing in the cardiac-muscle RNA as well as in the skeletal-muscle RNA, but inhibited the normal canonical splicing in the cardiac-

muscle RNA only. A related mechanism has been described in maize, where, as a consequence of a THE insertion/removal in the promoter region of the alcohol dehydrogenase 1 gene, it causes variations in the tissue-specific expression (Kloeckener-Gruissem and Freeling 1995). The role that the mobile elements actually play in causing changes in human DNA sequences and in gene expression deserves further investigation.

Acknowledgments

This work was entirely and generously financed by a grant from the British Heart Foundation (BHF) of Great Britain (to F.M.). A.F. is a BHF research fellow. We also wish to thank the "Legato Ferrari" Foundation (Modena, Italy) (support to A.F.); Dr. Michael Antoniou, for his support; and Dr. S. Abbs, for providing the DNA of a DMD patient with a large genomic deletion encompassing intron 11.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

EMBL, http://embl-heidelberg.de

GenBank, http://www.ncbi.nlm.nih.gov/Web/Genbank (for the rearranged 136-bp *Alu*-like sequence [accession number Y13186] and the wild-type 5′ end of intron 11, nucleotides 1–2371 [accession number Y13187])

Leiden Muscular Dystrophy Pages, http://www.dmd.nl

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for dystrophin gene [MIM 310200])

References

Adeniyi-Jones S, Zasloff M (1985) Transcription, processing and nuclear transport of a B1 Alu RNA species complementary to an intron of the murine α-fetoprotein gene. Nature 317:81–84

Ahn AH, Kunkel LM (1993) The structural and functional diversity of dystrophin. Nat Genet 3:283–291

Beggs AH, Koenig M, Boyce FM, Kunkel LM (1990) Detection of 98% of DMD/BMD gene deletions by polymerase chain reaction. Hum Genet 86:45–48

Britten RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. Proc Natl Acad Sci USA 93: 9374–9377

Britten RJ, Baron WF, Stout DB, Davidson EH (1988) Sources and evolution of human Alu repeated sequences. Proc Natl Acad Sci USA 85:4770–4774

Chamberlain JS, Gibbs RA, Ranier JE, Nguyen PN, Caskey CT (1988) Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. Nucleic Acids Res 16:11141–11156

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

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Craig NL (1995) Unity in transposition reactions. Science 270: 253–254

Deininger PL, Batzer MA, Hutchison CA, Edgell MH (1992) Master genes in mammalian repetitive DNA amplification. Trends Genet 8:307–311

Durand JB, Bachinski LL, Bieling LC, Czernusezewicz GZ, Abchee AB, Yu QT, Tapscott T, et al (1995) Localization of a gene responsible for familial dilated cardiomyopathy to chromosome 1q32. Circulation 92:3387–3391

Favor J, Morawetz C (1992) Insertional mutations in mammals and mammalian cells. Mutat Res 284:53–74

Ferlini A, Muntoni F (1998) The 5′ region of intron 11 of the dystrophin gene contains target sequences for mobile elements and three overlapping ORFs. Biochem Biophys Res Commun 242:401–406

Hahn SH, Krasnewich D, Brantly M, Kvittingen EA, Gahl WA (1995) Heterozygosity for an exon 12 splicing mutation and a W234G missense mutation in an American child with chronic tyrosinemia type 1. Hum Mutat 6:66–73

Holmes SE, Dombrosky BA, Krebs CM, Boehm CD, Kazazian HH (1994) A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. Nat Genet 7:143–148

Janicic N, Pausova Z, Cole DEC, Hendy GN (1995) 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

Kingsmore SF, Giros B, Suh D, Bieniarz M, Caron MG, Seldin MF (1994) Glycine receptor β-subunit gene mutation in spastic mouse associated with LINE-1 element insertion. Nat Genet 7:136–141

Kloeckener-Gruissem B, Freeling M (1995) Transposon-induced promoter scrambling: a mechanism for the evolution of new alleles. Proc Natl Acad Sci USA 92:1836–1840

Knebelmann B, Forestier L, Drouot L, Quinones S, Chuet C, Benessy F, Saus J et al (1995) Splice-mediated insertion of an Alu sequence in the COL4A3 mRNA causing autosomal recessive Alport syndrome. Hum Mol Genet 4:675–679

Krajinovic M, Pinamonti B, Sinagra G, Vatta M, Severini GM, Milasin J, Falaschi A, et al (1995) Linkage of familial dilated cardiomyopathy to chromosome 9. Am J Hum Genet 57: 846–852

Kunkel LM, Tantravahi U, Kurnit DM, Eisenhard M, Bruns GP, Latt SA (1983) Identification and isolation of transcribed human X chromosome DNA sequences. Nucleic Acids Res 11:7961–7979

Li L, Bray PF (1993) 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

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

Manolio TA, Baughman KL, Rodeheffer R, Pearson TA, Bristow JD, Michels VV, Abelmann WH, et al (1992) Prevalence and etiology of idiopathic dilated cardiomyopathy (summary of a National Heart, Lung and Blood Institute Workshop). Am J Cardiol 69:1458–1466

Mathias SL, Scott AF, Kazazian HH Jr, Boeke JD, Gabriel (1991) A reverse transcriptase encoded by a human transposable element. Science 254:1808–1810

McNaughton JC, Marshall CJ, Broom JE, Hughes G, Jones WA, Stockwell PA, Petersen GB (1995) Phylogenetic relationship among transposon-like elements in human and primate DNA. J Mol Evol 40:127–135

Milasin J, Muntoni F, Severini GM, Bartoloni L, Vatta M, Krjinovic M, Mateddu A, et al (1996) A point mutation in the 5′ splice site of the dystrophin gene first intron responsible for X-linked dilated cardiomyopathy. Hum Mol Genet 5:73–79

Mitchell GA, Labuda D, Fontaine G, Saudubray JM, Bonnefont JP, Lyonnet S, Brody LC, et al (1991) Splice-mediated insertion of an Alu sequence inactivates ornithine δ-aminotransferase: a role for Alu elements in human mutation. Proc Natl Acad Sci USA 88:815–819

Monaco AP, Bertelson CJ, Liechti-Gallati S, Moser H, Kunkel LM (1988) An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics 2:90–95

Muntoni F, Cau M, Ganau A, Congiu R, Arvedi G, Mateddu A, Marrosu MG, et al (1993) Deletion of the dystrophin muscle-promoter region associated with X-linked dilated cardiomyopathy. N Engl J Med 329:921–925

Muntoni F, Melis MA, Ganau A, Dubowitz V (1995) Transcription of the dystrophin gene in normal tissues and in skeletal muscle of a family with X-linked dilated cardiomyopathy. Am J Hum Genet 56:151–157

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

Narita N, Nishio H, Kitoh Y, Ishikawa Y, Ishikawa Y, Minami R, Nakamura H, et al (1993) 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 91:1862–1867

Nelson KK, Green MR (1990) Mechanism for cryptic splice site activation during pre-mRNA splicing. Proc Natl Acad Sci USA 87:6253–6257

Olson TM, Keating MT (1996) Mapping a cardiomyopathy locus to chromosome 3p22-p25. J Clin Invest 97:528–532

Pizzuti A, Pieretti M, Fenwick RG, Gibbs RA, Caskey CT (1992) A transposon-like element in the deletion-prone region of the dystrophin gene. Genomics 13:594–600

Richardson P, McKenna W, Bristow M, Maisch B, Mautner B, O'Connell J, Olsen E, et al (1996) Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of Cardiomyopathies. Circulation 93:841–842

Sebillon P, Beldjord C, Kaplan JC, Brody E, Marie J (1995) A T to G mutation in the polypyrimidine tract of the second intron of human beta-globin gene reduces in vitro splicing efficiency: evidence for an increased hnRNP C interaction. Nucleic Acids Res 23:3419–3425

Sewry CA, Sansome A, Clerk A, Sherratt TG, Hasson N, Rodillo E, Heckmatt JZ, et al (1993) Manifesting carriers of Xp21 muscular dystrophy: lack of correlation between dystrophin expression and clinical weakness. Neuromuscul Disord 3:141–148

Tanda S, Corces VG (1991) Retrotransposon-induced overexpression of a homeobox gene causes defects in eye morphogenesis in Drosophila. EMBO J 10:407–417

Towbin JA, Hejtmancik JF, Brink P, Gelb B, Zhu XM, Chamberlain JS, McCabe ER, et al (1993) X-linked cardiomyopathy: molecular genetic evidence of linkage to the Duchenne muscular dystrophy (dystrophin) gene at the Xp21 locus. Circulation 87:1854–1865

van der Kooi AJ, van Meegen M, Ledderhof TM, McNally EM, de Visser M, Bolhuis PA (1997) Genetic localization of a newly recognized autosomal dominant limb-girdle muscular dystrophy with cardiac involvement (LGMD1B) to chromosome 1q11-21. Am J Hum Genet 60:891–895

Vidaud D, Vidaud M, Bahnak BR, Siguret V, Sanchez SG, Laurian Y, Meyer D, et al (1993) 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

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

Weiner A, Deininger PL, Efstratiadis P (1986) A nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu Rev Biochem 55:631–661

Yoshida K, Ikeda S, Nakamura A, Kagoshima M, Takeda S, Shoji S, Yanagisawa N (1993) Molecular analysis of the Duchenne muscular dystrophy gene in patients with Becker muscular dystrophy presenting with dilated cardiomyopathy. Muscle Nerve 16:1161–1166