Hum Genet (2002) 110 : 418–421
DOI 10.1007/s00439-002-0721-7

ORIGINAL INVESTIGATION

**Carlo Nobile · Luisa Toffolatti · Francesca Rizzi
Barbara Simionati · Vincenzo Nigro · Barbara Cardazzo
Tomaso Patarnello · Giorgio Valle · Gian Antonio Danieli**

Analysis of 22 deletion breakpoints in dystrophin intron 49

Received: 29 October 2001 / Accepted: 25 February 2002 / Published online: 9 April 2002

Abstract Over 60% of Duchenne and Becker muscular dystrophies are caused by deletions spanning tens or hundreds of kilobases in the dystrophin gene. The molecular mechanisms underlying the loss of DNA at this genomic locus are not yet understood. By studying the distribution of deletion breakpoints at the genomic level, we have previously shown that intron 49 exhibits a higher relative density of breakpoints than most dystrophin introns. To determine whether the mechanisms leading to deletions in this intron preferentially involve specific sequence elements, we sublocalized 22 deletion endpoints along its length by a polymerase-chain-reaction-based approach and, in particular, analyzed the nucleotide sequences of five deletion junctions. Deletion breakpoints were homogeneously distributed throughout the intron length, and no extensive homology was observed between the sequences adjacent to each breakpoint. However, a short sequence able to curve the DNA molecule was found at or near three breakpoint junctions.

C. Nobile
CNR-Centro di Studio per la Biologia
e Fisiopatologia Muscolare,
c/o Dipartimento di Scienze Biomediche Sperimentali,
Università di Padova, 35121 Padua, Italy

L. Toffolatti · F. Rizzi · B. Cardazzo · T. Patarnello
G.A. Danieli (✉)
Dipartimento di Biologia, Università di Padova,
Viale G. Colombo 3, 35121 Padua, Italy
e-mail: danieli@civ.bio.unipd.it,
Tel.: +39-49-8276215Fax: +39-49-8276209

B. Simionati · G. Valle
CRIBI, Università di Padova, 35121 Padua, Italy

V. Nigro
Dipartimento di Patologia Generale,
Seconde Università di Napoli, 80138 Naples, Italy

T. Patarnello
Facoltà di Medicina Veterinaria "Agripolis",
Università di Padova, 35020 Padua, Italy

Introduction

Mutations causing Duchenne and Becker muscular dystrophies (DMD and BMD) occur in the dystrophin *(DMD)* gene, which is composed of 79 exons (Roberts et al. 1992) scattered over 2250 kb DNA (Nobile et al. 1997) at Xp21. The overall mutation rate in this gene is considerably higher than that in most X-linked and autosomal disease genes (Emery 1988). More than 60% of DMD and BMD mutations are deletions of variable sizes, affecting dystrophin exons with different frequencies. Higher deletion frequencies have been reported for two groups of contiguous exons in the 5' and central portions of the gene (Koenig et al. 1989; den Dunnen et al. 1989); these are referred to as the "minor" and "major" deletion hot spots, respectively. More recently, by analyzing the intragenic distribution of deletion breakpoints at the genomic level, we have shown that the increased breakpoint frequency near the 5' end of the gene directly correlates with the large sizes of some introns, whereas in the central intragenic region, the highest density of deletion breakpoints (number of breakpoints per intron length) occurs between exons 45 and 51 (Nobile et al. 1997). Precise mapping of deletion breakpoints along the lengths of introns 44 (Blonden et al. 1991) and 7 (McNaughton et al. 1998) has revealed no clustering of breakpoints, leading to the conclusion that deletion mechanisms operating in these introns are not mediated by specific sequences. Moreover, sequence analysis of nine deletion junctions in intron 7 have shown no extensive homology between the sequences adjacent to each breakpoint, ruling out unequal nonhomologous recombination as a major deletion mechanism in this intron (McNaughton et al. 1998). Both introns 7 and 44, however, lie outside the most mutable region of the dystrophin gene (Nobile et al. 1997). One of the smaller introns from this latter region, intron 49, exhibits a breakpoint density 4–5 times higher than those of introns 7 and 44 (see Koenig et al. 1989; http://www.dmd.nl/cdnagene.html). To check whether additional deletion mechanisms preferentially involving specific se-

Table 1 Intron-49 primer sequences (5'–3')

STS[a]	Forward primer	Reverse primer
A	CGATAAGTTAGCATCAAGCATTTG	CCTCAGAGCAGTTAATATGGTGAC
B	TTTTGAGCCTCTAGGTGTTTCAG	TTTGGAGAAAAATCAACATATCAAG
C	AAGAAGTTGACGAATGAAGGAAAG	ATTAAAGACACATCAGGGAGCAA
D	TATAATCTTCCAGGCATCATTTCC	TGTTGGCAGTAGAACAAACTCATT
E	AGCATCACATCTCTTGAGGTTATTT	TCTCTAAAATGCTTCCCTAATTTCA
F	GACTACAGTTTTTGATTTGGCTGTT	CAGCTTAATTGGAATGATGAGAGA
G	TGGGTATGTATGTGAAAGAAATGG	CTTTTGGGTTCAAGACTGATTTT
H	TCTTTTCCCCCAACTTGACTT	ATAAAATGATTGCACGCCTGG
I	GCTGCAAGAATGCAATTTATTAAG	TGTATCTCTGCACCCTTCTGATAA
J	TCCTATTGTGTTAAGCCACTGAAA	AACCCGACATCAATCAAGTTATTT

[a]*STS* Sequence tagged sites

quences and/or unequal crossing over account for the increased propensity to DNA breakage in intron 49, we sublocalized 22 DMD/BMD deletion breakpoints along this intron and analyzed the sequence of several deletion junctions.

Materials and methods

Genomic DNA (100 ng) from DMD or BMD patients was analyzed by the polymerase chain reaction (PCR) for 33 cycles (94°C for 1 min, 58°C for 1 min, and 72°C for 1 min). PCR products were fractionated by electrophoresis on a 2% agarose gel. Primer sequences specific for exons 49 and 50 have been reported previously (Beggs et al. 1990); intronic primer sequences are shown in Table 1.

Three breakpoints of deletions involving exons 48–49 were amplified by long-range PCR for 40 cycles (94°C for 15 s, 60°C for 1–8 min) by using rT*th* DNA Polymerase XL (PE Applied Biosystems). PCR products were resolved on a 1% agarose gel, purified, cloned into pCRII-TOPO vector (Invitrogen) and sequenced by using BigDye terminator chemistry (Perkin-Elmer/Applied Biosystems). Breakpoint junction sequences of deletions extending from intron 49 toward the 3' end of the gene were cloned by using the PCR-based Human GenomeWalker Kit (Clontech), following the manufacturer's instructions. Analysis of inverted repeats was performed by using Palindrome (http://www2.no.emb-

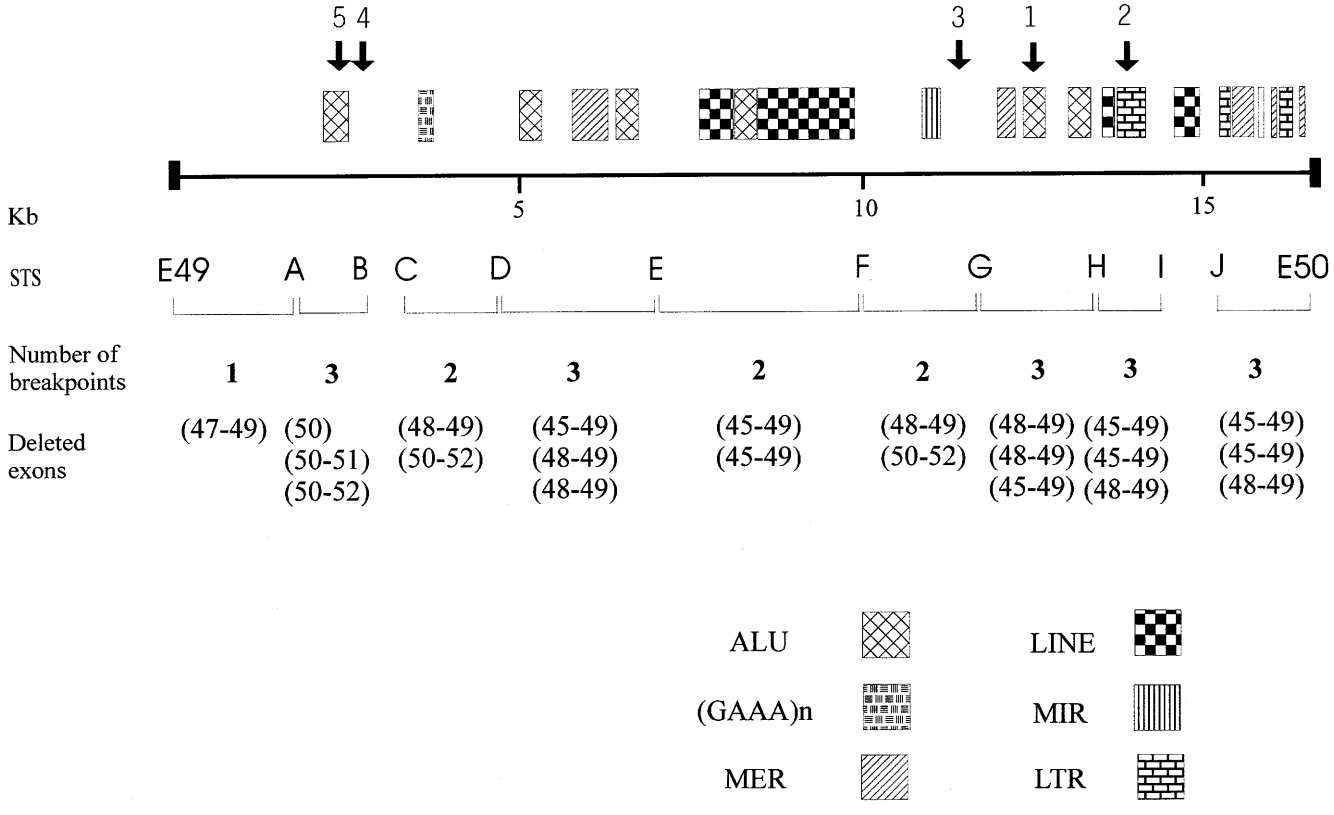

Fig. 1 Repetitive sequences and deletion breakpoint distribution in intron 49 of the human dystrophin gene. *Variously shaded boxes* Various classes of interspersed or simple repeat sequences, as indicated. *Arrows* Positions of the deletion junctions that were sequenced, *numbers above arrows* deletion junctions shown in Fig. 2. The positions of the intronic STSs used to localize deletion breakpoints are indicated by letters *A–J*. *Square brackets* Intronic intervals to which the indicated number of breakpoints were mapped. The missing exons of each deletion analyzed are indicated in *parentheses*

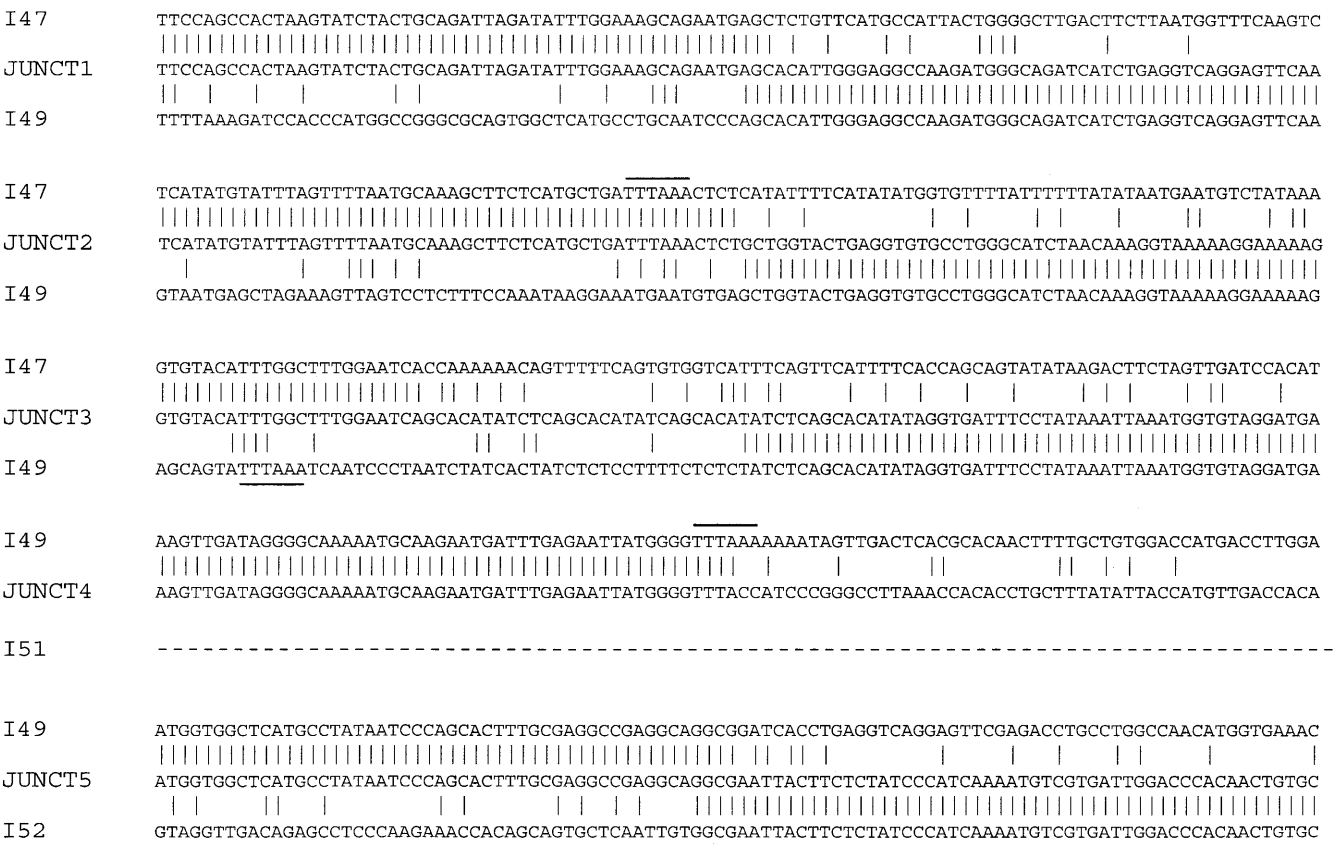

Fig. 2 DNA sequences spanning the five deletion breakpoints analyzed in detail, with the corresponding normal intron sequences, where available. Only the 5'–3' strands are shown. *Line* The sequence 5'-TTTAAA-3', where present

net.org/Programs/SAL/EMBOSS/). Possible matrix attachment regions (MAR) sites in intron 49 were identified by using MAR-WIZ (http://www.futuresoft.org/MAR-Wiz/).

Gene symbols used in this article follow the recommendations of the HUGO Gene Nomenclature Committee (Povey et al. 2001).

Results

The sequence of the dystrophin intron 49 was derived from genomic sequences released in public databases (accession nos. AC093193, AC021166, and AC079864), which allowed the assembly of a consensus sequence of 16,633 bp for this intron (accession no. AJ271220). Analysis of this intronic sequence performed by Repeatmasker (http://www.hgmp.mrc.ac.uk/GenomeWeb/Split/RepeatMasker) revealed a number of interspersed repeats, as schematically shown in Fig. 1. Based on sequence information, we designed PCR primers (A–J) for 10 sequence tagged sites (STSs; Table 1), dividing intron 49 into 11 contiguous intervals of roughly similar sizes (Fig. 1). The DNAs from 22 DMD or BMD patients, known to carry deletions with a breakpoint between exons 49 and 50, were tested for the presence or absence of each of these STSs to determine the intronic interval at which each deletion had its endpoint. A summary of the results of this analysis is shown in Fig. 1. In spite of the relatively small number of breakpoints analyzed, a uniform distribution of breakpoints is observed along the intron, with no particular intervals showing breakpoint clustering.

To PCR-amplify and analyze in more detail as many deletion junctions as possible, we employed two different approaches. Eight of the 22 deletions analyzed involved only exons 48 and 49 and therefore had the other (5') breakpoint in intron 47, which is about 55 kb long (http://www.dmd.nl/cdnagene.html). We were able approximately to sublocalize the 5'-breakpoints of these deletions in intron 47 by using the same STS-based approach described above (data not shown). Once a sufficiently small interval was defined to contain a given 5'-breakpoint, we attempted long-range PCR between the closest opposite primers flanking the deletion junction. In three cases, this strategy was successful, and the resulting PCR products were cloned and sequenced (sequences shown in Fig. 2; GenBank/DDBJ/EMBL accession nos. AJ430775, AJ430776, and AJ430777).

To PCR-amplify the breakpoint junctions of five deletions extending from intron 49 toward the 3'-end of the gene (see Fig. 1), we used a vectorette-like method. In two cases, we obtained PCR products across deletion junctions; these products were subsequently sequenced (results shown in Fig. 2; accession nos. AJ430778 and AJ430779). One of these junctions (no. 5) contains, on the 3'-side, a sequence from intron 52, whereas the other junction (no. 4), contains a 3'-sequence from intron 51

Table 2 Sequence types at five deletion junctions

Deletion junction	5' breakpoint		3' breakpoint	
	Intron	Sequence type	Intron	Sequence type
1	47	Unique	49	Alu
2	47	Unique	49	LTR
3	47	Unique	49	Unique
4	49	Unique	51	Unique
5	49	Alu	52	Unique

showing no homology with sequences currently released in genomic databases. As intron 51 has not completely been sequenced as yet, this latter sequence very likely derives from the intron-51 portion that is still unknown.

The positions of the five deletion breakpoints analyzed in detail with respect to the repetitive sequences present in intron 49 are shown in Fig. 1. Table 2 shows the types of sequence found at each side of these deletion junctions, and Fig. 2 shows the corresponding sequences, 50 bp on each side of the breakpoints. There is no extensive similarity between the DNA sequences adjacent to each breakpoint, and only a few basepairs of common DNA sequence are observed across the breakpoint junctions 1 and 4. Sequence features that might facilitate DNA breakage through exposure of single-stranded DNA, such as inverted repeated (palindromic) sequences, are not present at or near the breakpoints sequenced. The sequence TTTAAA, which is known to be able to curve the DNA molecule (Singh et al. 1997), is present at or in close proximity to junctions 2, 3, and 4. Junction 3, however, has more complex features: the sequence motif TCAGCACATA(TC), a single copy of which is present only in intron 49, just next to the breakpoint, is triplicated in the junction sequence, resulting in 27 bp extra DNA interposed between the joined intronic sequences. Finally, a search for MARs in intron 49 revealed that junction 1 occurs within a MAR.

Discussion

In this paper, we have shown that the remarkable increase of breakpoint density observed in dystrophin intron 49, with respect to introns 7 and 44, is not mediated by specific sequences particularly prone to DNA breakage, as indicated by the diffuse distribution of breakpoints along the intron. In addition, the lack of extensive homology between the sequences adjacent to the five breakpoints analyzed in detail does not support unequal nonhomologous recombination as a major mechanism for deletions in dystrophin intron 49, in agreement with previous studies performed on other, less mutable, regions of the dystrophin gene (McNaughton et al. 1998; Love et al. 1991; Brown et al. 1996). However, the sequence TTTAAA has been found at or in the vicinity of three breakpoints (Fig. 2); this sequence is known to be able to induce a curvature in the DNA molecule, which may predispose DNA to recombination and/or other nuclear processes (Singh et al 1997). Given the expected frequency of the sequence TTTAAA (1/1420 bp) in the human genome (Drmanac et al. 1986), the probability that three such sequences occurred by chance at or near the breakpoints is extremely low (3.49×10^{-10}), suggesting that curved DNA may be involved in one of the mechanisms leading to dystrophin deletions.

Acknowledgments The financial support of MURST is gratefully acknowledged.

References

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

Blonden LAJ, Grootscholten PM, Dunnen JT den, Bakker E, Abbs S, Bobrow M, Boehm C, Broeckoven C van, Baumback L, Chamberlain J, Caskey CT, Denton M, Felicetti L, Galluzzi G, Fishbeck KH, et al (1991) 242 breakpoints in the 200-kb deletion prone p20 region of the DMD gene are widely spread. Genomics 10:631–639

Brown J, Dry KL, Edgar AJ, Pryde FE, Hardwick LJ, Aldred MA, Lester DH, Boyle S, Kaplan J, Dufier JL, Ho MF, Monaco AP, Musarella MA, Wright AF (1996) Analysis of three deletion breakpoints in Xp21.1 and the further localization of RP3. Genomics 37:200–210

Drmanac R, Petrovic N, Glisin V, Crkvenjakov R (1986) A calculation of fragment lengths obtainable from human DNA with 78 restriction enzymes: an aid for cloning and mapping. Nucleic Acids Res 14:4691–4692

Dunnen JT den, Grootscholten PM, Bakker E, Blonden LAJ, Ginjar HB, Wapenaar MC, Passen HMB van, Broeckhonen C van, Pearson PL, Ommen GJB van (1989) Topography of Duchenne muscular dystrophy (DMD) gene: FIGE and cDNA analysis of 194 cases reveal 115 deletions and 13 duplications. Am J Hum Genet 45:835–847

Emery AEH (1988) Duchenne muscular dystrophy. Oxford Monographs on Medical Genetics, no. 15 (revised edn). Oxford University Press, New York

Koenig M, Beggs AH, Moyer M, Scherpf S, Heindrich K, Betteken T, Meng G, Muller CR, Lindlof M, Kaariainen H, Chapelle A de la, Kiuru A, Savontaus ML, Gilkenkrantz H, Recan D, et al (1989) The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. Am J Hum Genet 45:498–506

Love DR, England SB, Speer A, Marsden RF, Bloomfield JF, Roche AL, Cross GS, Mountford RC, Smith TJ, Davies KE (1991) Sequences of junction fragments in the deletion-prone region of the dystrophin gene. Genomics 10:57–67

McNaughton JC, Cockburn DJ, Hughes G, Jones WA, Laing NG, Ray PN, Stockwell PA, Petersen GB (1998) Is gene deletion in eukaryotes sequence-dependent? A study of nine deletion junctions and nineteen other deletion breakpoints in intron 7 of the human dystrophin gene. Gene 222:411–51

Nobile C, Marchi J, Nigro V, Roberts RG, Danieli GA (1997) Exon-intron organization of the human dystrophin gene. Genomics 45:421–424

Povey S, Lovering R, Bruford E, Wright M, Lush M, Wain H (2001) The HUGO Gene Nomenclature Committee (HGNC). Hum Genet DOI 101007/s00439-001-0615-0

Roberts RG, Coffey AG, Bobrow M, Bentley DR (1992) Determination of the exon structure of the distal portion of the dystrophin gene by vectorette PCR. Genomics 13:942–950

Singh GB, Kramer JA, Krawetz SA (1997) Mathematical model to predict regions of chromatin attachment to the nuclear matrix. Nucleic Acids Res 25:1419–1425