Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical and Biophysical Research Communications 347 (2006) 145–149
BBRC
www.elsevier.com/locate/ybbrc

A novel insertion of a rearranged L1 element in exon 44 of the dystrophin gene: Further evidence for possible bias in retroposon integration

Zuzana Musova [a], Petra Hedvicakova [a], Marketa Mohrmann [a], Marketa Tesarova [b], Anna Krepelova [a], Jiri Zeman [b], Zdenek Sedlacek [a,*]

[a] *Institute of Biology and Medical Genetics, Charles University Second Medical School and University Hospital Motol, Prague, Czech Republic*
[b] *Department of Paediatrics and Center of Applied Genomics, Charles University First Medical School, Prague, Czech Republic*

Received 5 June 2006
Available online 21 June 2006

Abstract

L1 elements are mammalian retrotransposons contributing to genome evolution and causing rare mutations in human. We describe a *de novo* insertion of an L1 element into the dystrophin gene resulting in skipping of exon 44 and causing Duchenne muscular dystrophy in a boy. The L1 element was rearranged due to the twin-priming mechanism, but contrary to all described L1 rearrangements the 5′ region of the inverted L1 sequence ended within the poly(A) tail of the element. Furthermore, the target site for the insertion was located only 87 bp from the insertion site in another patient described previously. These findings can contribute to the understanding of the mechanisms of L1 element rearrangement, and may support the notion that some subregions of the human genome could be preferred targets for retroelements using the L1 enzymatic machinery.

Keywords: L1 retrotransposon; Twin-priming; Disease-causing insertion; Duchenne muscular dystrophy; Exon skipping; Insertion bias

L1 elements are mammalian autonomous retrotransposons capable of being transcribed into RNA, and reverse transcribed and integrated back into a new chromosomal site. L1 elements contain two open reading frames (ORF1 and ORF2) coding for proteins needed for their retrotransposition (an RNA-binding protein and endonuclease/reverse transcriptase). The enzymatic machinery of the L1 elements contributes also to retroposition of non-autonomous retroposons Alu or SVA, and processed pseudogenes [1,2]. Several dozens of potentially active full-length L1 elements of about 6 kb [3] and several hundred thousands of truncated, rearranged or otherwise mutated L1 copies comprise almost 20% of the human genome [4]. The retrotransposition of L1 elements is expected to be a germline event but rare instances of somatic insertion in cancer have also been described [5]. Since the first description in 1988 of haemophilia caused by L1 insertions into the *F8* gene [6], a total of 15 human disease-causing mutations involving new L1 retrotransposition have been described [1,7]. Although *in vitro* assays and analyses of chromosomal distribution of young L1 elements indicate that their integration into the genome may be random or near-random [8,9], the compilation of human disease-causing insertions suggests several levels of possible insertion bias: preference of the X chromosome for L1 insertion, multiple insertions into nearby chromosomal sites, and even occurrence of independent insertions of elements using the L1 machinery into exactly the same site [7,10]. In this report, we describe a *de novo* insertion of an L1 element into the dystrophin gene causing Duchenne muscular dystrophy (OMIM 310200) in an 8-year-old boy. The current insertion site was very close to a site of another disease-causing L1 insertion reported previously [11], thus supporting the possibility of L1 insertion bias.

* Corresponding author. Fax: +420 224433520.
E-mail address: zdenek.sedlacek@lfmotol.cuni.cz (Z. Sedlacek).

0006-291X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2006.06.071

Materials and methods

Patient. The boy was born from an uncomplicated pregnancy with birth weight 4100 g and length 53 cm. Postnatal adaptation and motor development within the first two years of life were uneventful with some delay in speech and social behaviour. At the age of 18 months, enlargement of calves was observed. Laboratory analyses showed increased levels of creatine kinase (CK, 194–324 μkat/l; controls <2.5 μkat/l), myoglobin (629–1261 μg/l; controls <80 μg/l), and aminotransferases (ALT 7.5 μkat/l; AST 5.97 μkat/l). The Gowers' sign was evident by the age of 4 years. A psychological test at the age of 6 years revealed subnormal intellect. At present at the age of 8 years the disease shows a mild progression. Both parents and two brothers of the boy are healthy with normal levels of CK. Informed consent was obtained prior to the genetic analysis.

Mutation analysis. The analysis of deletions in the dystrophin gene was performed on genomic DNA using a modification of the standard multiplex PCR method [12,13]. Exons were amplified in several sets, and those not amplified in the multiplex reactions were amplified once more with other controls to confirm the deletion. The critical exon 44 was amplified using primers 44F (GTTGTGTGTACATGCTAGGTGTGTA) and 44R (TCCATCACCCTTCAGAACCTGATCT). Total blood lymphocyte RNA was isolated using TRI Reagent (Molecular Research Center) and reverse transcribed into cDNA using oligo(dT) primer and SuperScript III RNAse H-Reverse Transcriptase according to the protocol of the manufacturer (Invitrogen). One microliter of the reaction was amplified in a semi-nested reaction by primers specific for exons 43 and 45 of the dystrophin cDNA (44FC (TTGCAAAGTGCAACGCCTGTGG), 44RC (TGACAGCTGTTTGCAGACCTCC), and 44RintC (TCTGACAACA GTTTGCCGCTGC)). PCR products were analysed on 2.5% agarose gels. Sequencing was performed with the BigDye Terminator v3.1 Cycle Sequencing Kit, ABI 3100 sequencer (Applied Biosystems), and primers used in PCR or internal L1-specific primer L1R1 (ATGGCTAGCC AGTTTTCCCAGC).

Results

Characterisation of the insertion on the genomic DNA level

The analysis of the dystrophin gene of the patient resulted in the identification of a *de novo* L1 insertion in exon 44 (Fig. 1). The first line of evidence for an insertion came from the observation of a longer PCR product after the amplification of this exon from genomic DNA of the patient, while both his mother and a healthy brother as well as an unrelated normal control individual yielded a fragment of the expected size (Fig. 2A). Sequencing of the aberrant fragment showed an L1 element insertion between codons 2108 and 2109 of the *Dp427c* isoform of the dystrophin gene (Fig. 1). The poly(A) tail of the L1 element indicated insertion in the antisense orientation with respect to the dystrophin gene transcription. The sequence of the insertion site was 5′-TT/AAAT-3′, similar to the published consensus (5′-TT/AAAA-3′) [14]. The insertion was flanked by direct repeats (target site duplication) of 16 bp. The L1 element itself was rearranged. The polyA tail of 41 bp was preceded by 411 bp of sequence corresponding to bases 4735–5145 of the L1 Ta consensus [15], but in reverse orientation (Fig. 1). This sequence was 100% identical to the L1RP element (GenBank Accession No. AF148856) [16], and differed from the L1 Ta consensus by a C to T change at position 5131 of the consensus. The element thus belonged most likely to the Ta-1 (Ta-1d) subset of L1 sequences [15].

Influence of the insertion on the dystrophin gene expression

In silico analysis suggested that the insertion may have theoretically preserved the open reading frame of the dystrophin gene creating a protein with insertion of a tract of 14 phenylalanine residues followed by a stretch of 137 amino acid residues of the reverse transcriptase (encoded by ORF2 of the L1 element) and another five residues encoded by the target site duplication (Fig. 1). The expression of the dystrophin gene was tested by RT-PCR on RNA transcripts from lymphocytes of the patient and by sequencing of the RT-PCR products. A major novel transcript was identified with exon 44 skipped, and exon 43 joined directly to exon 45 (Fig. 2B and C). As exon 44 is 148 bp long, its skipping created a frameshift and generated a STOP codon 16 triplets downstream of the novel exon–exon junction.

Discussion

Mechanisms of the L1 element rearrangement and dystrophin gene silencing

We present a new L1 insertion in the dystrophin gene leading to exon skipping and Duchenne muscular dystrophy in the patient. Similarly to the vast majority of disease-producing L1 elements, the present L1 sequence belonged to the Ta subset [1,7]. The inserted element was truncated and rearranged due to the twin-priming mechanism [17]. While one reverse transcription started from the 3′ end of the L1 RNA, a second internal priming occurred upstream by annealing of the second end of the broken target DNA. Indeed, there was a significant homology between the last bases of the target site duplication and the internal region of the L1 sequence (four of the last five bases, Fig. 1), consistent with internal priming. Also, a microhomology of 1 bp at the junction of the two juxtaposed L1 segments was present (Fig. 1), consistent with subsequent annealing of the two cDNA single-strands. The inversion point was located at nucleotide 5145 of the L1 Ta consensus, within the common inversion hot-spot [17]. However, the current insertion was unique as the 5′ region of the inverted L1 segment ended within the poly(A) tail of the element. This structure of the rearrangement indicated that the L1 reverse transcriptase most likely dissociated from the stretch primed from the poly(T) primer before it reached the internal primer. This scenario was proposed to be less common compared to dissociation after reaching the internal primer [17], but it could explain why the length of the current poly(A) tail (41 bp) was shorter than the average length of 60–90 nucleotides observed in other disease-causing L1 insertions [7]. The short length of the poly(A) tail and the much larger size of the inverted segment (411 bp) could also support another possible feature of the twin-priming mechanism, "that invasion of the internal primer onto the L1 RNA has already occurred either before reverse transcription has begun at the poly(T)

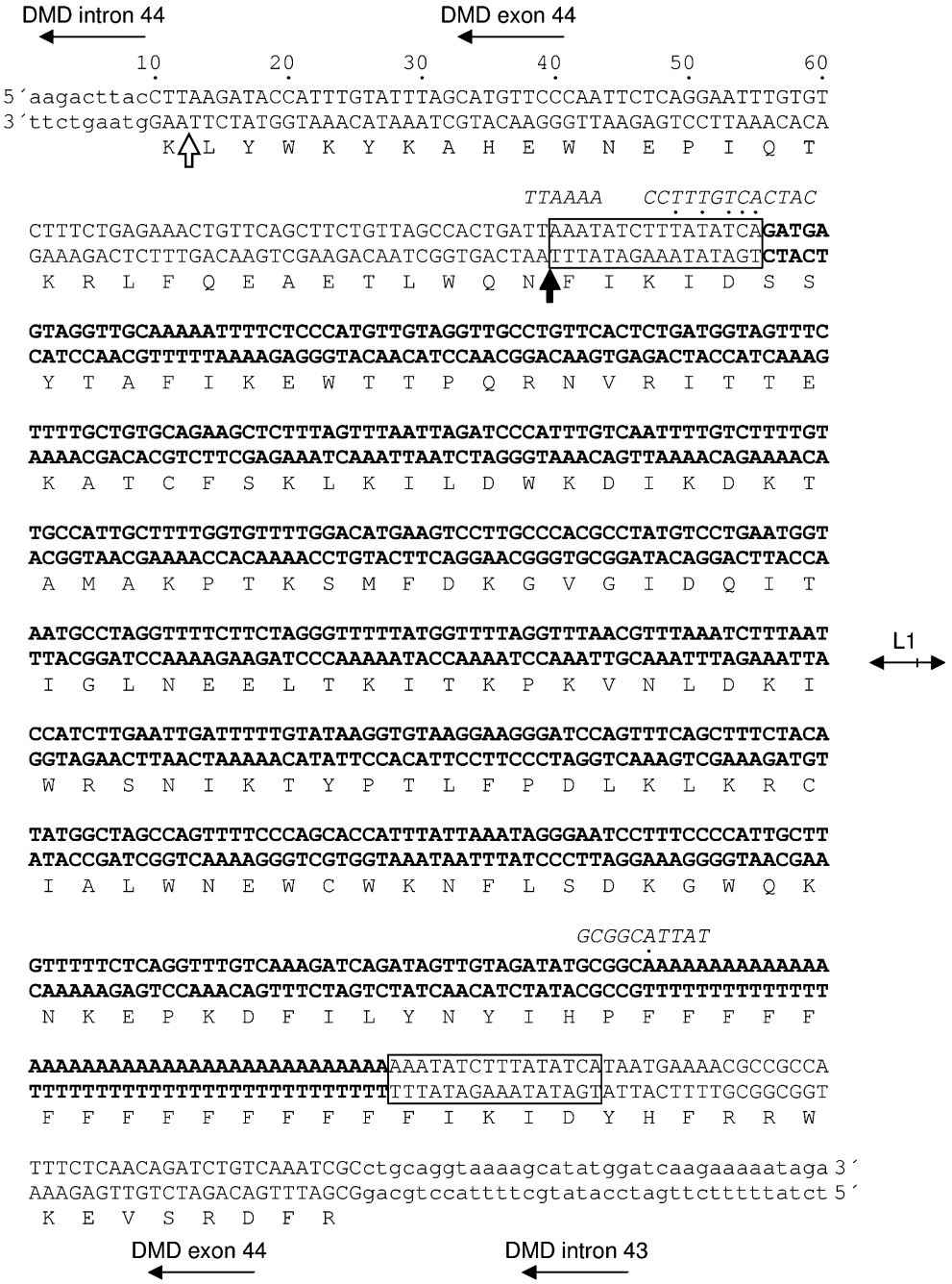

Fig. 1. Sequence of dystrophin exon 44 with the L1 insertion. Double-stranded DNA sequence of exon 44 (capital letters) and parts of introns 43 and 44 (small letters) is presented in the antisense orientation. The L1 insertion (bold) and the target site duplications (boxed) are indicated as well as the *in silico* translation (of the bottom strand), the insertion site consensus TTAAAA (italics), and two fragments of sequence of the L1 element (italics) with homology used in twin-priming (dots, see text for details). Close distance of the insertion site in the current patient (black arrow) and a case described previously [11] (open arrow) is shown.

primer, or shortly thereafter" [17]. Although a theoretical possibility existed that the insertion event could have preserved the reading frame and produced a protein with a stretch of illegitimate amino acid residues, the cellular splicing machinery produced an mRNA with the critical exon skipped as the major product. This was similar to other disease-causing L1 insertions which very often led to various splicing defects [11,18].

Possible sources of L1 insertion bias

The proportion of human mutations caused by retroposition was estimated at about 0.2% [1,2,19]. The example of dystrophin where exon deletions are routinely screened on genomic DNA based on the absence of the respective PCR products illustrates that this may be an underestimate. Long insertions may not be amplifiable and may be

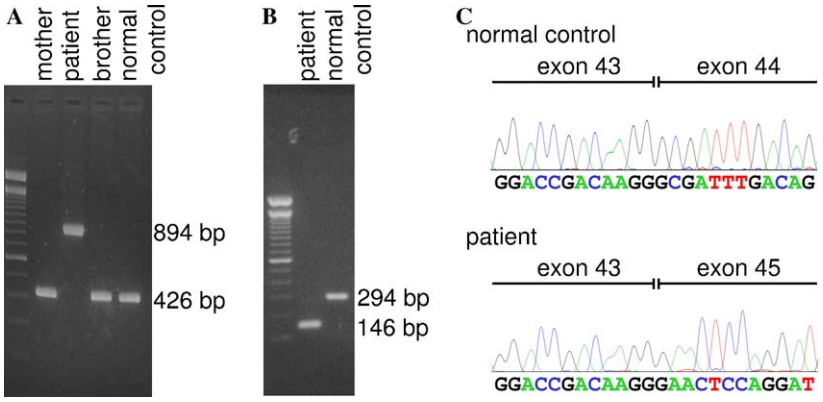

Fig. 2. Molecular genetic analysis of the insertion. (A) Amplification of exon 44 from genomic DNA of the patient, his relatives, and a normal control using flanking intronic primers. The longer fragment indicates an insertion. DNA size marker: 100 bp DNA ladder (Invitrogen). (B) Amplification of exon 44 from blood lymphocyte cDNA of the patient and a normal control using exon 43 and exon 45 primers. The shorter fragment suggests skipping of exon 44 in the patient. (C) Sequence analysis of the cDNA fragments from (B) yields evidence on the nucleotide level for skipping of exon 44.

misinterpreted (in the absence of a confirmatory test like Southern blot, MLPA, long-range PCR of junction fragments, etc.) as deletions. Similarly, intronic insertions affecting splicing are unlikely to be identified by methods using screening of exons on the genomic DNA level. Finally, PCR-based methods cannot identify large insertions in carriers of X-linked disorders or at autosomal loci, where the insertions are masked by the normal-sized product from the second allele.

This detection bias could account for higher frequency of retroposon insertion mutagenesis in human X-linked diseases. It is interesting, however, that while disease-causing L1 insertions are found almost exclusively on the X chromosome, the partition between the X chromosome and autosomes is almost equal for disease-causing Alu insertions, and opposite for SVA insertions [1,7]. A general preference of L1 elements for the X chromosome cannot be used to explain this difference. Although the X chromosome has a higher overall L1 content [4], the young L1 Ta-1 subfamily, which is responsible for most of the disease-causing insertions, does not show any bias for the sex chromosomes [9]. However, excess recruitment of functional retrogenes caused partly by insertion bias was observed on the X chromosome [20]. It is clear that separation of detection bias on the one hand from insertion bias and post-insertion selection on the other hand is very difficult [21,22].

In addition to possible bias described above, several interesting instances of independent integration events of retroelements using the L1 machinery into two very close sites or even into the same site, which were unlikely to occur by chance, have been described. These include insertions into orthologous sites of different species or strains [23], but mainly independent human disease-causing insertions observed in unrelated patients [7,10]. Two L1 insertions within 615 bp of the *F8* gene [6], two Alu insertions within 105 bp of the *FGFR2* gene [24], an L1 and an Alu insertion into the same site of the *APC* gene [5,25], two Alu insertions into the same site of the *F9* gene [26,27], and an Alu and a SVA insertion into the same site of the *BTK* gene [10] have been described. The L1 insertion described here may be another such example. The insertion site is located only 87 bp proximal from the L1 insertion site described in another patient [11]. The selection of target sites by L1 sequences and other retroelements using the L1 machinery may be influenced by the chromatin state [28] or possibly by broader sequence context, higher chromosome structure, nuclear localization, involvement of the segment in cellular processes like replication, transcription and repair and their timing, and other potential factors which remain to be determined. In addition to *in vitro* analyses, reporting of novel disease-causing integrations can significantly contribute to this aim because they represent authentic *in vivo* events, and unlike fixed or polymorphic elements they were affected much less by selection on the population level.

Acknowledgment

This work was supported by Grant MZO00064203 from the Ministry of Health of the Czech Republic.

References

[1] E.M. Ostertag, H.H. Kazazian Jr., Biology of mammalian L1 retrotransposons, Annu. Rev. Genet. 35 (2001) 501–538.
[2] H.H. Kazazian Jr., Mobile elements: drivers of genome evolution, Science 303 (2004) 1626–1632.
[3] B. Brouha, J. Schustak, R.M. Badge, S. Lutz-Prigge, A.H. Farley, J.V. Moran, H.H. Kazazian Jr., Hot L1s account for the bulk of retrotransposition in the human population, Proc. Natl. Acad. Sci. USA 100 (2003) 5280–5285.
[4] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, et al., Initial sequencing and analysis of the human genome, Nature 409 (2001) 860–921.
[5] Y. Miki, I. Nishisho, A. Horii, Y. Miyoshi, J. Utsunomiya, K.W. Kinzler, B. Vogelstein, Y. Nakamura, Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer, Cancer Res. 52 (1992) 643–645.
[6] H.H. Kazazian Jr., C. Wong, H. Youssoufian, A.F. Scott, D.G. Phillips, S.E. Antonarakis, Haemophilia A resulting from de novo

insertion of L1 sequences represents a novel mechanism for mutation in man, Nature 332 (1988) 164–166.
[7] J.M. Chen, P.D. Stenson, D.N. Cooper, C. Ferec, A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease, Hum. Genet. 117 (2005) 411–427.
[8] D.V. Babushok, E.M. Ostertag, C.E. Courtney, J.M. Choi, H.H. Kazazian Jr., L1 integration in a transgenic mouse model, Genome Res. 16 (2006) 240–250.
[9] S. Boissinot, A. Entezam, L. Young, P.J. Munson, A.V. Furano, The insertional history of an active family of L1 retrotransposons in humans, Genome Res. 14 (2004) 1221–1231.
[10] M.E. Conley, J.D. Partain, S.M. Norland, S.A. Shurtleff, H.H. Kazazian Jr., Two independent retrotransposon insertions at the same site within the coding region of BTK, Hum. Mutat. 25 (2005) 324–325.
[11] N. Narita, H. Nishio, Y. Kitoh, Y. Ishikawa, Y. Ishikawa, R. Minami, H. Nakamura, M. Matsuo, 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 (1993) 1862–1867.
[12] J.S. Chamberlain, R.A. Gibbs, J.E. Ranier, C.T. Caskey, Multiplex PCR for the diagnosis of Duchenne muscular dystrophy, in: M.A. Innis, D.H. Gelfand, J.J. Sninsky, T.J. White (Eds.), PCR Protocols: A Guide To Methods and Applications, Academic Press, San Diego, 1990, pp. 272–281.
[13] S. Abbs, S.C. Yau, S. Clark, C.G. Mathew, M. Bobrow, A convenient multiplex PCR system for the detection of dystrophin gene deletions: a comparative analysis with cDNA hybridisation shows mistypings by both methods, J. Med. Genet. 28 (1991) 304–311.
[14] J. Jurka, Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons, Proc. Natl. Acad. Sci. USA 94 (1997) 1872–1877.
[15] S. Boissinot, P. Chevret, A.V. Furano, L1 (LINE-1) retrotransposon evolution and amplification in recent human history, Mol. Biol. Evol. 17 (2000) 915–928.
[16] M.L. Kimberland, V. Divoky, J. Prchal, U. Schwahn, W. Berger, H.H. Kazazian Jr., Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells, Hum. Mol. Genet. 8 (1999) 1557–1560.
[17] E.M. Ostertag, H.H. Kazazian Jr., Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition, Genome Res. 11 (2001) 2059–2065.
[18] C. Meischl, M. Boer, A. Ahlin, D. Roos, A new exon created by intronic insertion of a rearranged LINE-1 element as the cause of chronic granulomatous disease, Eur. J. Hum. Genet. 8 (2000) 697–703.
[19] X. Li, W.A. Scaringe, K.A. Hill, S. Roberts, A. Mengos, D. Careri, M.T. Pinto, C.K. Kasper, S.S. Sommer, Frequency of recent retrotransposition events in the human factor IX gene, Hum. Mutat. 17 (2001) 511–519.
[20] J.J. Emerson, H. Kaessmann, E. Betran, M. Long, Extensive gene traffic on the mammalian X chromosome, Science 303 (2004) 537–540.
[21] J.M. Chen, C. Ferec, D.N. Cooper, LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease: mutation detection bias and multiple mechanisms of target gene disruption, J. Biomed. Biotechnol. 2006 (2006) Article ID 56182, 9pp.
[22] T. Graham, S. Boissinot, The genomic distribution of L1 elements: the role of insertion bias and natural selection, J. Biomed. Biotechnol. 2006 (2006) Article ID 75327, 5pp.
[23] H.J. Ho, D.A. Ray, A.H. Salem, J.S. Myers, M.A. Batzer, Straightening out the LINEs: LINE-1 orthologous loci, Genomics 85 (2005) 201–207.
[24] M. Oldridge, E.H. Zackai, D.M. McDonald-McGinn, S. Iseki, G.M. Morriss-Kay, S.R. Twigg, D. Johnson, S.A. Wall, W. Jiang, C. Theda, E.W. Jabs, A.O. Wilkie, De novo alu-element insertions in FGFR2 identify a distinct pathological basis for Apert syndrome, Am. J. Hum. Genet. 64 (1999) 446–461.
[25] K.C. Halling, C.R. Lazzaro, R. Honchel, J.A. Bufill, S.M. Powell, C.A. Arndt, N.M. Lindor, Hereditary desmoid disease in a family with a germline Alu I repeat mutation of the APC gene, Hum. Hered. 49 (1999) 97–102.
[26] D. Vidaud, M. Vidaud, B.R. Bahnak, V. Siguret, S. Gispert Sanchez, Y. Laurian, D. Meyer, M. Goossens, J.M. Lavergne, 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 (1993) 30–36.
[27] K. Wulff, H. Gazda, W. Schroder, R. Robicka-Milewska, F.H. Herrmann, Identification of a novel large F9 gene mutation-an insertion of an Alu repeated DNA element in exon e of the factor 9 gene, Hum. Mutat. 15 (2000) 299.
[28] G.J. Cost, A. Golding, M.S. Schlissel, J.D. Boeke, Target DNA chromatinization modulates nicking by L1 endonuclease, Nucleic Acids Res. 29 (2001) 573–577.