Journal of Human Genetics (2012) 57, 92–100

www.nature.com/jhg

ORIGINAL ARTICLE

CAPN3 mRNA processing alteration caused by splicing mutation associated with novel genomic rearrangement of *Alu* elements

Ikhlass Hadj Salem[1], Ines Hsairi[2], Najla Mezghani[1], Houda Kenoun[1], Chahnez Triki[2] and Faiza Fakhfakh[1]

Recessive mutations of *CAPN3* gene are reported to be responsible for limb girdle muscular dystrophy type 2A (LGMD2A). In all, 15–25% of intronic nucleotide changes identified in this gene were investigated by *in silico* analysis, but occasionally supported by experimental data or reported in some cases as a polymorphism. We report here genetic and transcriptional analyses in three Tunisian patients belonging to the same consanguineous family sharing the same mutation c.1194-9 A>G and *Alu* repeats insertion in intron 7 of *CAPN3* gene. Reverse transcriptase-PCR experiments performed on total RNA from the patient's muscle biopsy showed retention of the eight last nucleotides of intron 9 in the *CAPN3* transcript lacking the first seven exons. Our results provide evidence regarding the potential involvement of *Alu* elements in aberrant processing of pre-mRNA owing to the disruption of pre-existing intronic splicing regulatory elements. We also demonstrated variable mRNA alternative splicing among tissues and between LGMD2A patients. A deep intronic variation and rearrangement have been reported in the literature as causing genetic diseases in humans. However, this is the first report on a potential pathogenic *CAPN3* gene mutation resulting from an *Alu* insertion.
Journal of Human Genetics (2012) **57,** 92–100; doi:10.1038/jhg.2011.129; published online 8 December 2011

Keywords: *Alu* element; alternative splicing; CAPN3; intronic variant; LGMD2A

INTRODUCTION

Mutations in *CAPN3* gene (MIM no. 114240) cause an autosomal recessive form of limb girdle muscular dystrophy type 2A (LGMD2A, MIM no. 253600).[1,2] *CAPN3* gene encodes for a non-structural protein, the enzyme called calpain-3 (originally named p94).[3] Calpain-3 is the muscle-specific member of a family of Ca^{2+}-dependent proteases, which are supposed to have a role in many intracellular processes, including cell motility, apoptosis, differentiation and cell cycle regulation, by modulating the biological activity of their substrates through limited and strictly controlled proteolysis.[4] Calpain-3 is composed of four functional domains: domain PC1 (protease core domain 1) has a regulatory role, domain PC2 (protease core domain 2) is the proteolytic module, domain C2L (C2 domain-like) has a C2-domain-like Ca^{2+} binding function and domain penta-EF-hand binds Ca^{2+} ions.[5,6]

LGMD2A is one of the most frequently occurring forms of LGMD, which is characterized by a very high genetic variability.[7–10] Predominant symmetrical and simultaneous involvement of pelvic and scapular girdle and trunk muscles without facial, oculo-motor or cardiac involvements is typical for LGMD2A.[11–13] Over 440 distinct *CAPN3* gene mutations have so far been reported on the Leiden Muscular Dystrophy Database as the cause of a wide spectrum of clinical severity.[14] Approximately 70% of mutant alleles are of missense type; the remainders are null mutations (deletion/insertion causing frame shifting, nonsense and splice site mutations), large genomic rearrangements and synonymous or intronic changes causing aberrant splicing.[15] Many CAPN3 intronic variants have been identified during diagnostic screening: they account for about 15% of all variants listed on the Leiden Database, and for about 25% of the mutations reported in other studies.[15–17] For the majority of intronic variants the consequences on mRNA splicing have been only inferred by *in-silico* analysis, whereas experimental demonstration of their pathogenicity has been obtained by mRNA studies for only 1% of them (Leiden Database).[15–18] Although deep intronic sequences were originally believed to be non-functional, because they do not code for proteins, it has been suggested that some of these sequences do indeed have relevance.[19] In most cases, deep intronic disease-causing variations or rearrangements could affect gene splicing directly by disruption of pre-existing intronic splicing-regulatory elements.[20–23]

In this study, we performed a molecular genetic analysis of the *CAPN3* gene in three female siblings of a consanguineous family with the LGMD2A phenotype. The availability of diagnostic muscle biopsy

[1]Laboratoire de Génétique Moléculaire Humaine, Faculté de Médecine de Sfax, Université de Sfax, Sfax, Tunisia and [2]Service de Neurologie Pédiatrique, C.H.U. Hédi Chaker de Sfax, Sfax, Tunisia
Correspondence: Dr IH Salem or Professor F Fakhfakh, Laboratoire de Génétique Moléculaire Humaine, Faculté de Médecine de Sfax, Université de Sfax, Avenue Magida Boulila, 3029 Sfax, Tunisia.
E-mails: ikhlasshadj@gmail.com or faiza.fakhfakh@gnet.tn
Received 28 August 2011; revised 19 October 2011; accepted 21 October 2011; published online 8 December 2011

and RNA splicing analyses from blood samples served to demonstrate pathogenetic effects, at both transcriptional and translational levels, of the c.1194-9A>G splicing mutation and *Alu* repeats insertion in the *CAPN3* gene.

MATERIALS AND METHODS

Three affected females of a Tunisian family were born to consanguineous parents. The age of the affected individuals ranged between 9 and 20 years. Four additional healthy siblings and the parents were also recruited. Informed consent was obtained from patients and control individuals in accordance with the ethics committee of the University Hospital of Sfax. Affected members underwent general neurological examinations. Clinical history and physical examinations of the family members ruled out the implication of environmental factors in LGMD etiology. Blood samples were collected from eight family members and fifty healthy individuals. Genomic DNA was extracted from whole blood following a standard phenol–chloroform method.[24]

Microsatellite genotyping and mutation analysis

For each gene and locus responsible for LGMD2 and previously described in Tunisian LGMD2 patients (*CAPN3*, *DYSF*, *SGCA*, *SGCB*, *SGCG* and *FKRP*), at least two microsatellite markers were selected on the basis of their map position and heterozygosity coefficient. Fluorescent dye-labeled microsatellite markers were genotyped for all the family members. We used the True Allele PCR Premix (Applied Biosystems, Foster City, CA, USA) for PCR reactions according to the manufacturer's instructions. Fluorescently labelled alleles were analysed on an ABI PRISM 3100-Avant automated Genetic Analyser (Applied Biosystems). Genotypes were determined using the GenScan software (Applied Biosystems).

One affected subject was investigated for the presence of mutation in *CAPN3* gene as previously described.[25] The amplified products of all coding exons, exon–intron junctions and promoter region were directly sequenced using an ABI 3100-Avant automated DNA sequencer and Big Dye Terminator Sequencing V3.1 Kit (Applied Biosystems).

Bioinformatics prediction of splice consensus score and protein mutant model

Online promoter analysis tools of TESS at http://www.cbil.upenn.edu/cgi-bin/tess/tess[26] and TRANSFAC at http://www.gene-regulation.com/pub/databases.html#transfac[27] were used to search for the possible sequence of the transcriptional binding site and to assess the potential functional significance of −665C>G (no. rs3098421). To evaluate the strength of the altered splice-site of c.1194-9 A>G mutation (DB-ID CAPN3_00088), splice site scores were predicted by the following splice-site prediction programs (SSPPs): Splice-Site Prediction by Neural Network (NNSPLICE V0.9 at http://www.fruitfly.org/seq_tools/splice.html),[28] Human Splicing Finder (HSF V2.4 at http://www.umd.be/HSF/)[29] and Splicescan II at (http://splicescan2.lumc.edu/).[30] The mutation genomic sequence environment was also analyzed using the online ESEfinder web interface (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home).[31]

RNA extraction and reverse transcriptase (RT)-PCR analysis

Total RNA from the skeletal muscle of patients P1 and P2 was obtained using the SV Total RNA Isolation System (Promega, Madison, WI, USA) including treatment with DNaseI. CAPN3 cDNA (based on NM_000070) was amplified into eight overlapping regions. The fragment of cDNA containing c.1194-9 A>G mutation was amplified using primers designed in exon 8 (5′-GGTGGAGTGGAACGG-3′) and exon 11 (5′-CTCCGAGTCATCAGGG-3′) and was expected to yield a 340-bp fragment. An independent control cDNA amplification of a 368-bp fragment from exons 3 to 7 of the *SGCG6* gene using the primer pair 5′-CTCAGAAGGGGAGGTCACAG-3′ and 5′-CAGCATCAAGCACAAGCATT-3′ was done in order to test the quality of cDNA templates used for the analysis. The amplification of CAPN3 and SGCG cDNA was done in three independent replicates.

The functional effect of the found variations c.1194-9A>G and −665C>G was also assessed by RT-PCR analysis of lymphoid RNAs obtained from affected and control individuals. Total RNA was isolated from 10 ml of blood samples using PureLink Micro-to-Midi Total RNA Purification System (Invitrogen, Karlsruhe, Germany). RT-PCR, covering the CAPN3 cDNA, was carried out for the three patients using the same primers used to detect mRNA changes of the *CAPN3* gene in muscle, according to the manufacturer's recommendations of SuperScript Tm One-Step RT-PCR with platinum Taq kit (Invitrogen). Direct sequencing of RT-PCR products was performed by standard conditions in both directions.

As normal controls, we used cDNA from muscle tissue and blood sample from subjects who were free of any neuromuscular disorders. Nucleic acids were quantified using the Nano Drop ND-1000 UV–Vis spectrophotometer (Thermo SCIENTIFIC, Wilmington, DE, USA).

Immunoblot analysis

Semi-quantitative analysis of control, LGMD2C and P2 patients' muscle homogenates was performed with calpain-3 monoclonal antibody (NCL-Calp3c/12A2 against exon 8; Novocastra, Newcastle, UK) as previously described.[25] Myosin heavy-chain staining with Coomassie blue on the post-blotted gel was used as a protein-loading control.

Amplification and characterization of the CAPN3 rearrangement

Long-range PCR was performed on genomic DNA from the three probands to detect intronic rearrangement. The PCR reaction was carried out in a 20-µl reaction volume using the Long PCR Enzyme Mix (Fermentas, Burlington, Ontario, Canada) according to the manufacturer's instructions, with primers located in the flanking exon 7 and within intron 7 (F: 5′-GTAAGCC TGGTGGGGCTTGGTG-3′; R: 5′-CTTAAGCACACAGAAAGAGC-3′) of the *CAPN3* gene. The following PCR cycles were used for amplification: initial denaturation at 94 °C for 3 min, 10 cycles at 94 °C for 15 s, 61 °C for 30 s, 68 °C for 8 min, followed by 24 cycles under the same conditions with elongation of the extension step for 20 s in each cycle, and a final extension at 68 °C for 20 min. The PCR product was loaded on a 1% agarose gel and visualized by ethidium bromide staining.

The specificity of the PCR product was verified by direct sequencing of the extremities. Sequences were analyzed using the Blast 2 Sequences program at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) and RepeatMasker software (http://www.repeatmasker.org/).

RESULTS

We identified a family with three affected female siblings. The two elder sisters presented a progressive myopathic syndrome, whereas the younger one was apparently normal. The healthy parents were first-degree cousins and originated from southern Tunisia.

Clinical variability of LGMD2A within a single family

Proband 1. Proband 1 was first assessed in the Child Neurology department at Hedi Chaker Hospital in Sfax at the age of 18 years. She had normal motor milestone development, and fatigability in walking and running appeared at the age of 17 years. At the age of 20 years, physical examination revealed an intermediate clinical course showing positive Gowers sign and a waddling gait without equinus foot. Muscle strength of proximal muscles of the upper and lower limbs was estimated at 4 (according to the Medical Research Council (MRC) Scale of Muscle Strength). Deep tendon reflexes were normal. The patient also presented scapular winging and slight hyperlordosis without calves hypertrophy. The plasma creatine phosphokinase level was elevated, at 2272 UIl^{-1} (normal range $\leqslant$350 IUl^{-1}). Her electromyogram was significant for diffuse myopathic changes.

Proband 2. The sister of proband 1, 14 years old, was referred to the Child Neurology department at Hedi Chaker Hospital in Sfax at the age of 12 years for progressive difficulties in walking. Her medical history revealed normal motor milestone development and a muscle weakness starting at the age of 10 years leading progressively to

fatigability in walking and running. Physical examination showed a severe clinical course, relative to that of her sister, showing a waddling gait with equinus foot with advanced-stage symptoms, such as difficulty in climbing stairs, rising up from a chair or getting up from the floor without help. Muscle weakness involved the proximal muscles of the upper and lower limbs, with the muscle strength estimated at 3. There was no evidence of distal involvement. Muscle tone was low, and deep tendon reflexes were weak. The patient presented also scapular winging, thigh amyotrophy with calves hypertrophy and severe hyperlordosis. The plasma creatine phosphokinase level was elevated, at 5198 UI l^{-1}. Her electromyogram was significant for diffuse myopathic changes.

Proband 3. The younger sister of probands 1 and 2 was 9 years old and was systematically examined. Physical examination revealed normal walking, no muscle weakness and obtainable deep tendon reflex. The plasma creatine phosphokinase level was elevated, at 6876 UI l^{-1}.

A point mutation within *CAPN3* gene revealed by nucleotide sequencing

In an attempt to identify the responsible gene, we first performed a genetic linkage analysis of fluorescent dye-labeled polymorphic microsatellite markers covering all LGMD2 loci described so far in Tunisia. This analysis revealed evidence for linkage to CAPN3 mapping to chromosome 15q. All affected individuals showed a homozygous haplotype for alleles 210, 184 and 193 bp of D15S514, D15S781 and D15S222 microsatellite markers, respectively (Figure 1a). Direct sequencing of the exon–intron junctions and promoter region of *CAPN3* gene revealed no mutations in the 24 exons that could affect the function of calpain-3. However, direct sequencing in the three probands revealed a known C to G transversion (−665C>G) in the homozygous state within an *Alu* element located in the promoter region and an A to G transition (c.1194-9 A>G) in the homozygous state located in position −9 upstream of the invariable intronic AG motif of the acceptor splice site adjacent to exon 10. Cosegregation of c.1194-9 A>G and −665C>G variations with LGMD2A in the family members was also determined by direct sequencing (Figure 1b).

The single-nucleotide mutation enhances partial retention of intron 9 by creating an additional upstream acceptor site

To investigate the possible effect of −665C>G and c.1194-9A>G on mRNA transcriptional level and/or splicing, we performed bioinformatics and RT-PCR analyses. In fact, the bioinformatics analysis using Transcription Element Search Software (TESS) indicated that −665C>G transversion (no. rs3098421) might alter transcript factor binding such that the G allele created a potential binding site (CTNGTG) for a transcriptional repressor HES-1 (Hairy and Enhancer of Split 1) (Figure 2a). The c.1194-9A>G mutation occurred at position −9 upstream of exon 10. An *in silico* analysis predicted that this mutation potentially interfered with the correct splicing, making more probable a different acceptor splice site localized 9 bases downstream from the canonical one. The new acceptor site increased the predicted scores from 87, 51 in the wild-type allele to 91, 77 in the mutated allele and caused the insertion of the last eight nucleotides of intron 9 in mature RNA (Figure 2a). The ESEfinder program showed that the c.1194-9 A>G substitution was predicted to create an exon-splicing enhancer in intron 9 of the *CAPN3* gene. The new exon-splicing enhancer was TTCCAGC, which was predicted to be recognized by SRp40 SR protein with a score of 2.77, and is thus higher than the threshold value of 2.67 (Figure 2b).

RT-PCR analysis was performed on *CAPN3* mRNA obtained from the muscle biopsy of patients P1 and P2. Using the primers within exons 8 and 11, the control sample displayed an expected band of 340 bp. In contrast, a slightly longer band appeared in LGMD2A patient P1 (Figures 3a and b). The sequence of the mutant band showed retention of the last eight nucleotides of intron 9 in *CAPN3* transcript (r.1148_1149ins1149-8_1149-1) (Figure 3f). However, no amplification product was found for proband P2 although we tested the quality of cDNA by amplification of a 368-bp control fragment of the *SGCG* gene showing amplification in the cDNA of control patients, P1 and P2 (Figure 3d).

Unexpectedly, amplification of the entire *CAPN3* coding region (exons 1–24) in overlapping RT-PCR reactions using RNA isolated from muscle biopsy of a healthy individual and LGMD2A patients showed that the new alternatively spliced form of CAPN3 mRNA muscle lacked the first seven exons (exons 1–7) in LGMD2A patient P1 (Figures 3b and c).

In view of the fact that no tissue sample was available for mRNA analysis as regards patient P3, and to further investigate a new alternative approach for performing the molecular diagnosis, we examined calpain 3 expression in white blood cells (WBCs) to perform the molecular diagnosis of LGMD2A at mRNA level. In contrast to CAPN3 expression examined in muscle, we were able to amplify and sequence all designed *CAPN3* fragments on cDNA from the blood of a healthy control individual and LGMD2A patients. In the three LGMD2A patients, blood cDNA showed similar results for the retention caused by c.1194-9 A>G (Figure 3e). Results of western blot analysis performed for patient P2 using calpain-3 antibody are shown in Figure 3g. Normal protein levels for calpain-3 were observed in normal and LGMD2C muscle biopsies used as a control, whereas a total deficiency in calpain-3 was found in patient P2 (absence of the 94- and 60-kDa bands in the blot) (Figure 3g). No muscle samples were available for protein analysis for patients P1 and P3.

Novel insertion of *Alu* elements into the *CAPN3* gene could alter splicing sites within intron 7

Based on the result obtained in the RNA analysis and explaining the deletion of the first seven exons, the computer analysis showed that intron 7 (total length: 1533 nucleotides) of *CAPN3* gene naturally harbored *Alu* sequences between nucleotides 861 and 1170. This repeat element showed sequence homology with the Sg subfamily of *Alu* sequence. We used long-range PCR to amplify a fragment within intron 7 of the *CAPN3* gene from the patient's genomic DNA. A 2.5-kb PCR product was obtained instead of the expected 1.5-kb product (Figure 4). The extremities of the abnormal PCR product were sequenced and sequence alignment with the normal genomic sequence of CAPN3 revealed that there was an insertion of an equivalent of three *Alu* elements residing head to head in close proximity within intron 7 in these probands, as there was at least 1 kb difference compared with the normal PCR size.

DISCUSSION

In the present work, we report three sisters with LGMD2 phenotype associated with homozygosity for more than one potential pathogenic variant including a splicing mutation c. 1194-9 A>G, *Alu* repeats insertion in intron 7 of *CAPN3* gene and the −665C>G known variation. The first variant, c-1194-9 A nucleotide, lay in position 9 upstream of the invariable intronic AG motif of the acceptor splice site adjacent to exon 10. This variation was identified in several previous reports (Table 1), where its pathogenetic effects have been suggested.[7,32,33] Indeed, because of the lack of a molecular proof,

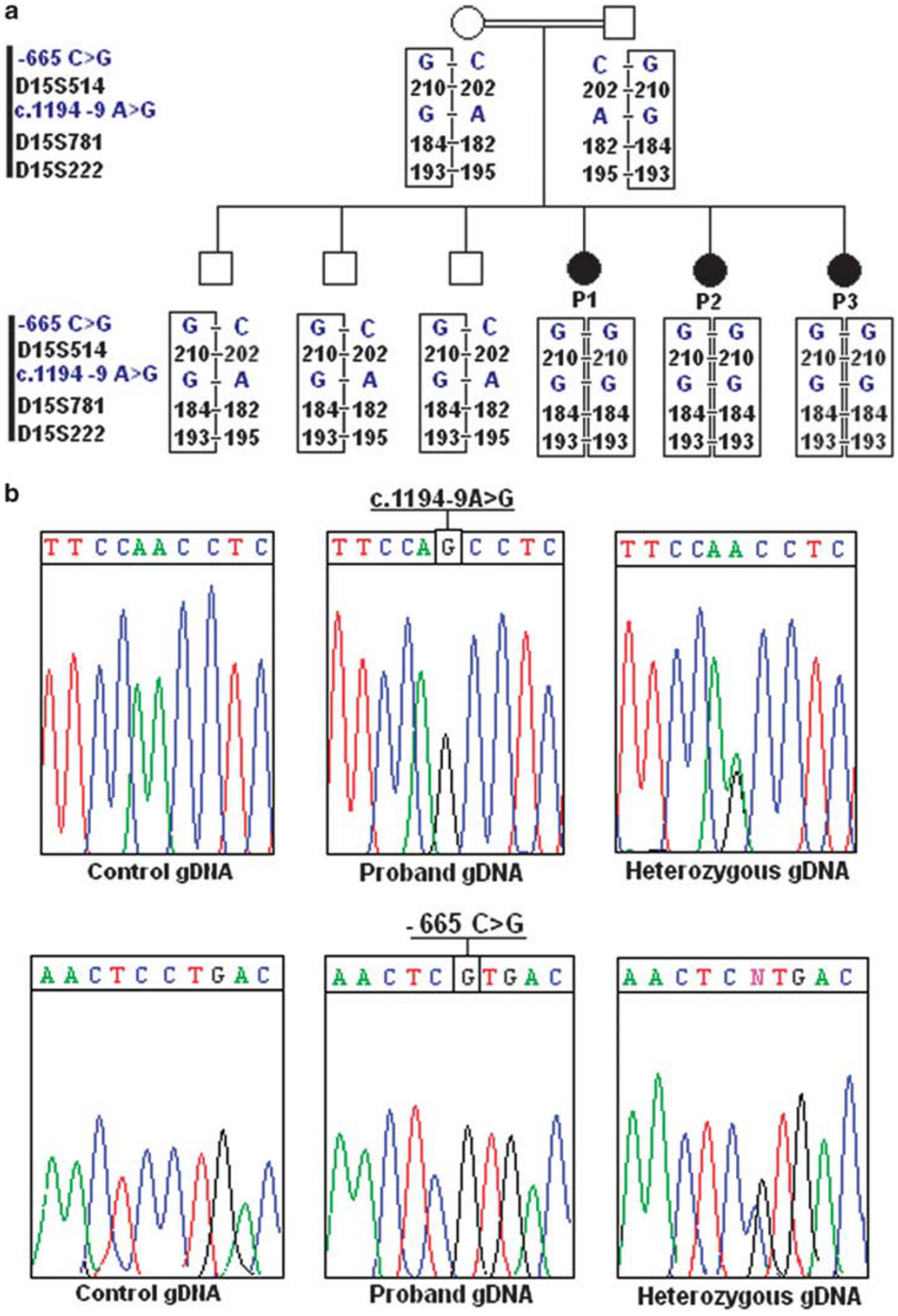

Figure 1 *CAPN3* gene linkage and mutation analyses. (**a**) Pedigree of the Tunisian family showing the segregation of CAPN3 haplotype and the inheritance of the c.1194-9 A>G mutation and −665C>G polymorphism. The haplotypes containing the mutation are boxed. (**b**) Direct genomic DNA (gDNA) sequence from a healthy control subject, the proband P1 and one heterozygous subject. The proband was homozygous for a single-nucleotide change in her genomic DNA G>A at the intron 9 near the consensus splice donor site.

this mutation as well as some intronic variants have been reported either as polymorphisms or as 'possibly pathogenetic', thus generating confusion and compromising a conclusive genetic counselling. None of these studies have provided a functional effect of the c.1194-9 A>G mutation. In our study, computer-assisted analyses suggested that G>A exchange disrupted the recognition of the acceptor splice site adjacent to exon 10 and created a new cryptic splice site in the region surrounding this mutation, causing the retention of the last 8 nucleotides of intron 9. Its deleterious effect was definitely demonstrated at a transcriptional level in our study. Indeed, RNA analysis revealed the insertion of the last eight nucleotides of intron 9 in mature RNA. The analysis by the online ESEfinder database of normal and variant sequences in intron 9 suggested a possible creation of an enhancer, which could be responsible for the synthesis of an alternative isoform of *CAPN3* mRNA. This idea has already been reported in HLA-G transcript[34] and in the mRNA encoding the β-catalytic subunit of the mitochondrial H+ATP synthase, where translation-enhancing activity generates a new isoform.[35]

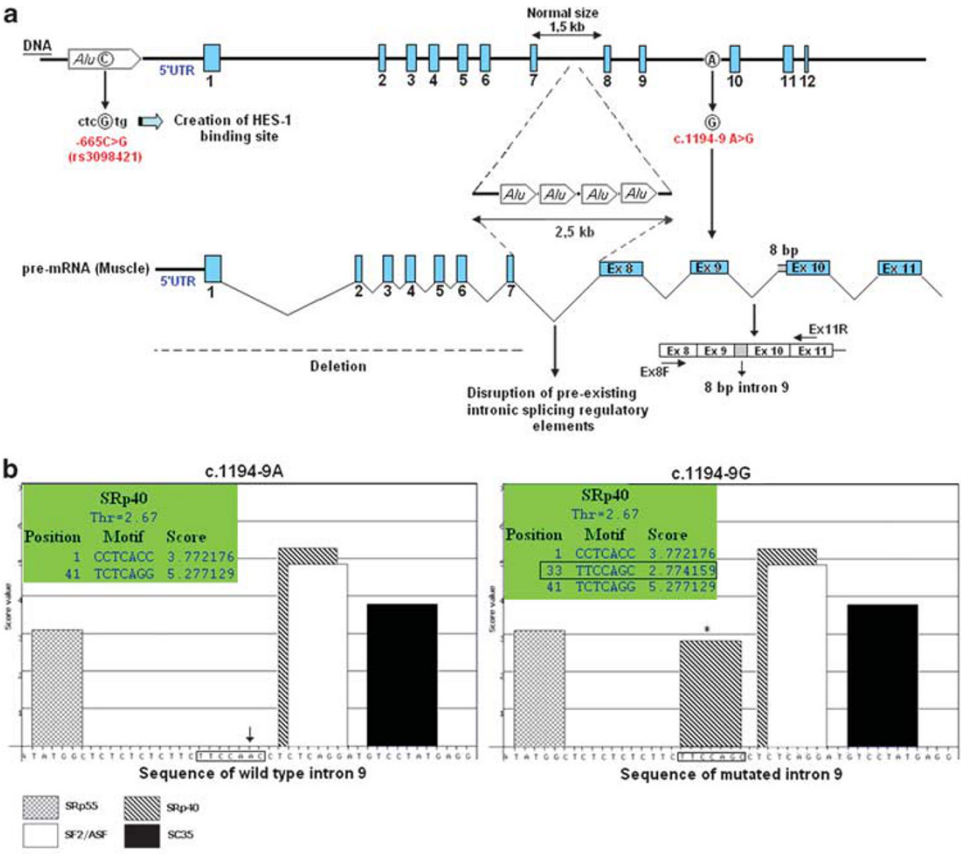

Figure 2 Detected DNA changes and *in silico* prediction. (**a**) Schematic representation of the bioinformatics prediction of c.1194-9 A>G splicing mutation. Patients had an A>G substitution at position -9, upstream of exon 10 associated to a multi insertion event of *Alu* elements in intron 7. This mutation causes an additional splicing of the mRNA transcript, with incorporation of eight nucleotides from intron 9 in a transcript lacking the seven first exons. *Alu* insertion event is predicted to be responsible for the disruption of pre-existing intronic splicing regulatory elements. (**b**) Computational prediction of the effect of c.1194-9 A>G splicing mutation change on enhancer composition. The A at position −9 upstream the exon 10 of *CAPN3* gene does not include an ESE. Note that the A to G substitution creates a new enhancer in intron 9 of *CAPN3* gene recognized by SRp40 SR protein. The arrows indicate the substitution position.

Many studies have examined how apparently similar mutations can have different effects on how the pre-mRNA is processed and how variability in mRNA splicing can potentially affect the resulting phenotype. This is the case of variation c.1746-20C>G in intron 13 of *CAPN3* gene. Indeed, the splicing analysis of the c.1746-20 C>G showed five splicing patterns resulting from the use of new cryptic acceptor sites, carrying either different intronic portions or the entire intron 13. These transcripts are variably expressed and not always detectable in the mutant patients.[15,18,36] In our study, it seems reasonable to deduce that the observed variability between two siblings sharing the same mutation is owing to the ability to splice the mutant allele. The clinical heterogeneity observed in this family supported findings of previous studies concerning CAPN3-splicing mutations (Table 1). Krahn *et al.*[17] reported on two patients who carried the mutation c.802-9 G>A in *CAPN3* gene at a heterozygous state. The effect of the c.802-9 G>A variant differed between the patients, which could be partially caused by inter-individual variation in NMD efficiency related to yet unexplained genetic modifying factors.

In our study, conversely to CAPN3 mRNA obtained from the biopsy sample, when CAPN3 expression was examined in white blood cells, it showed a constant expression both in patients and control, suggesting that mRNA splicing was achieved through several mechanisms, which are usually tissue-specific to assure the fidelity of gene expression.[16] Our study supported the alternative approach of performing the molecular diagnosis by using mRNA from peripheral blood as p94-calpain isoform was found to be expressed in human circulating peripheral blood mononuclear cells.[37]

Unexpectedly, we also observed the existence of an alternatively spliced form of CAPN3 mRNA with a missing region spanning from exon 1 to exon 7. Theoretical effect of this deletion, at both transcriptional and translational levels, is difficult to be investigated, as *CAPN3* gene was found to contain more than one promoter region, which might result independently in the production of new isoform of calpain-3.[38] We speculate that this deletion could be related to non-identified variation within either intron 7 or upstream exon 1 leading to complete absence of the first part of mRNA. In our assay, we showed that amplification product for intron 7 from the genomic DNA of the probands was larger (≈2.5 kb) than the expected size (≈1.5 kb) and revealed after sequencing in both directions an *Alu* multi insertion event lying in intron 7 of *CAPN3* gene. The intron 7 sequence was modified after *Alu* insertion, which might affect intron processing directly by creating or strengthening a splice site leading to an alternative splicing in the final transcript. The impact it might have on gene splicing upon integration in introns is still to be investigated.

In fact, *Alu* repeats are short interspersed elements whose transposition has been repeatedly implicated in genetic variability and

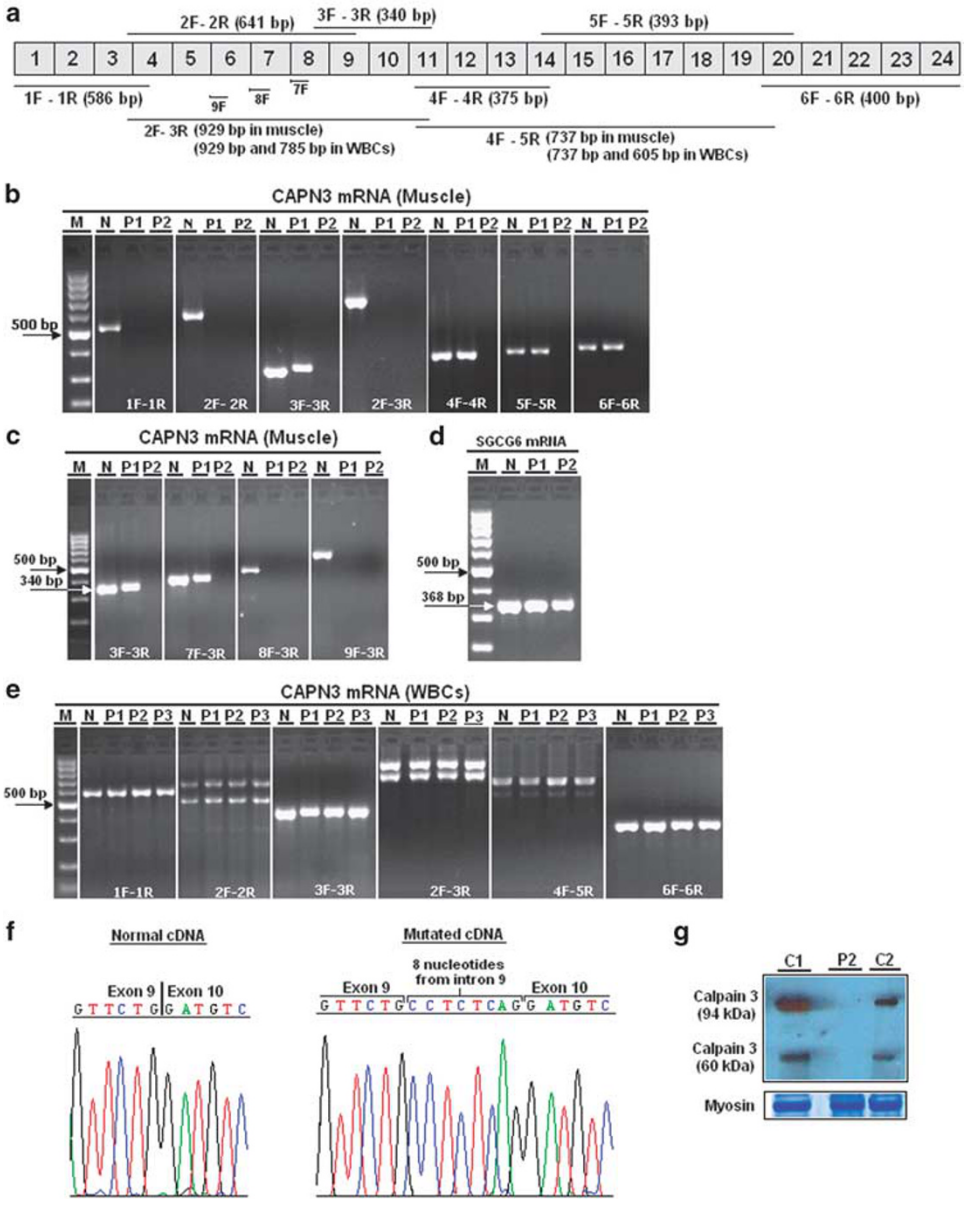

Figure 3 Detection of messenger RNA (mRNA) changes by RT-PCR analysis and calpain3 deficiency by western blot. (**a**) Schematic representation of the primers used for the amplification of CAPN3 cDNA; (**b**, **c**) gel electrophoresis of RT-PCR amplification product of all fragments in muscle: the region spanning exons 8 to11 (3F–3R) of *CAPN3* gene in muscle showing: a single band in the normal at the expected size of 340 bp (N) and a slightly longer one of 348 bp in proband P1, whereas no amplification was found for proband P2. No amplification was found for region spanning exons 1 to 7 for patient P1. (**d**) Amplification of a 368 bp control fragment from exons 3–7 of *SGCG6* gene on cDNA from normal subject (N) and patients P1 and P2. (**e**) Gel electrophoresis of the RT-PCR amplification product of all fragments of the *CAPN3* gene in WBCs showing a constant expression in patients P1, P2, P3 and in control (N). (**f**) Sequence chromatograms of amplified cDNA products in the normal subject and patient P1. Note the partial retention of the last eight nucleotides of intron 9. The size marker is a 100-bp DNA ladder (M). (**g**) Detection of calpain3 deficiency by western blot using monoclonal antibodies NCL–CALP–12A2. Calpain-3 western blotting in muscle biopsy of control (C1) and LGMD2C patient (C2) showed normal level of calpain-3, whereas a total deficiency of the 94- and 60-kDa bands was observed in patient P2 (P2). An amount comparable to the normal control, as determined by myosin in the post-transfer Coomassie blue-stained gel.

heritable disorders mainly by providing the site for homologous recombination[39] and alternative splicing.[40–42] *Alu* insertion/deletion polymorphism has been reported to be in total linkage disequilibrium with CTG repeats in myotonic dystrophy.[43] Several reports[44–49] indicate that *de novo Alu* insertions into intronic sequences in close proximity to the affected exon cause the downstream exon to shift from constitutive splicing to full exon skipping or alternative splicing (Table 2). In our study, this event occurred in a tissue-specific manner, suggesting rather a disruption of pre-existing intronic splicing regulatory elements. Indeed, increasing evidence shows that long introns contain several potential splicing regulatory sequences, including cryptic splice sites and splicing enhancers or silencers that, when activated, can be involved in aberrant processing of pre-mRNA. Moreover, the *Alu* consensus contains 9 potential 5′ splice sites and 14 potential 3′ splice sites.[50] The presence of several potential splice sites in the *Alu* consensus sequence[50,51] strongly suggested that they

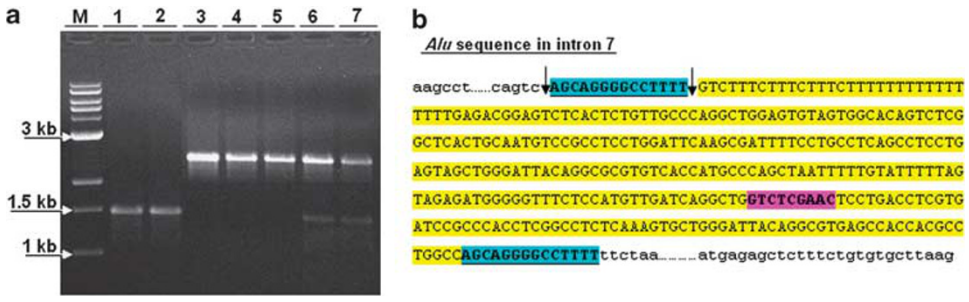

Figure 4 Detected *Alu* elements insertion in intron 7 of *CAPN3* gene. (**a**) Gel electrophoresis of the long-range PCR amplification product of intron 7 of *CAPN3* gene. Lanes 1 and 2 are normal controls, lanes 3, 4 and 5 are LGMD2A patient samples P1, P2 and P3, respectively, lane 6 is the patient's father and lane 7 is the patient's mother. A 2.5-kb PCR product was obtained instead of the expected 1.5-kb product. Lane M is the size marker of the 1-kb DNA ladder. (**b**) Intron 7 of *CAPN3* gene with sequence elements of the *Alu* repeat (*Alu* repeat is highlighted in yellow). Consensus target site is shown in red bold characters; target-site duplications of *CAPN3* gene sequence flanking the integrated DNA are highlighted in sky blue. The arrows indicate the cuts.

Table 1 Summary of cases carrying the c.1194-9 A>G mutation and other splicing mutations in *CAPN3* gene

Origin	*Splicing mutation*	*Gender*	*State*	*Age at onset (years)*	*Age at ascertainment*	*Calpain deficiency*	*Transcriptional and mRNA study*	*Reference*
Brazilian	c.1194-9 A>G	F	Homo	14	25	Partial	NP	32
Japanese	c.1194-9 A>G	F	Homo	20	48	NP	NP	33
UK	c.1194-9 A>G	F	Hetero with c.2263+1 G>C	ND	ND	NP	NP	7
French	c.1194-9 A>G	ND	Hetero	ND	ND	NP	NP	Leiden Database
Tunisian	c.1194-9 A>G	F	Homo	10	Still walking	Complete	Retention of 8 bp from intron 9	Present study
Tunisian	c.1194-9 A>G	F	Homo	17	Still walking	NP	Retention of 8 bp from intron 9	Present study
Veneto region (Italy)	c.1193+6T>A	F	Hetero	ND	ND	Partial	Retention of 31 bp from intron 9	15
Spain	c.802-9 G>A	F	Hetero	12	Still walking	Complete	Retention of 7 bp from intron 5	16
French	c.802-9 G>A	ND	Hetero with c.1714C>T	ND	ND	Severe decrease	Undetectable allele	17
French	c.802-9 G>A	ND	Hetero with c.2380+12delA	13	ND	Complete	Pseudo-homozygous insertion of 7 bp from intron 5	17
Tunisian	c.1536+1G>T	F/M	Homo	[7–16][a]	[24–33][a]	Complete	Total retention of intron 12	25

Abbreviations: F, female; Hetero, heterozygous; Homo, homozygous; M, male; ND, no current data; NP, not performed.
[a]Denotes ages at onset and at ascertainment ranged between 7–16 and 24–33 years, respectively.

Table 2 Diseases resulting from *Alu* insertion within an intron[41]

Gene	*Disease*	*Alu*	*Intron insertion*	*Orientation*	*Distance from SS*	*Effect*	*Reference*
CTDP1	CCFDN syndrome	Yf4	6	Antisense	73 bp upstream of exon 6	Alternative splicing	44
GK	Glycerol kinase deficiency	Y	4	Antisense	52 bp upstream of exon 5	Alternative splicing	45
FGFR2	Apert syndrome	Ya5	8	Antisense	19 bp upstream of exon 9	Alternative splicing	46
NF1	Neurofibromatosis type1	Ya5	5	Antisense	44 bp upstream of exon 6	Skipping of exon 6	47
FVIII	Hemophilia A X-linked severe bleeding disorder	Yb9	18	Antisense	19 bp upstream of exon 19	Skipping of exon 19	48
Fas	ALPS syndrome	Sb1	7	Antisense	50 bp upstream of exon 8	Skipping of exon 8	49

Abbreviations: ALPS, autoimmune lymphoproliferative syndrome; CCFDN, congenital cataracts, facial dysmorphin and neuropathy.

were recruited in the coding region through exonization. Of the thousands of *Alu* elements that are found in introns of the human genome, a certain number of complete or partial *Alu* sequences are also present in the coding regions of mature mRNAs.[52]

Additionally, the mutational analysis revealed a previously described promoter polymorphism, −665C>G (no. rs3098421) in *CAPN3* gene. *In silico* prediction showed that this C to G transversion created HES-1 binding site (CTNGTG) between −668 and −663 nucleotides. HES-1, Hairy and Enhancer of Split homolog-1, belongs to a family of basic helix-loop-helix transcriptional repressor that has an essential role in several developmental processes including myogenesis, neurogenesis, hematopoiesis and sex determinant.[53] HES-1 either antagonizes positive basic helix–loop–helix transcription factors through the E box (CANNTG) or directly binds to the N box (CTNGTG) and recruits a complex with co-repressors.[54] These data suggest the

possible involvement of HES-1 in CAPN3 transcription efficacity in muscle cells.

In this study, more than one mechanism seemed to be involved in generating the pathogenic phenotype of LGMD2A. We established that the c.1194-9A>G mutation identified in a Tunisian family with LGMD2A resulted in the partial retention of eight nucleotide of intron 9. Although this mutation was expected for a long time as a polymorphism, we provided a definite demonstration of its pathogenetic effect. We described also an *Alu*Sg insertion event in intron 7 of *CAPN3* gene, which could lead to the disruption of pre-existing intronic-splicing regulatory elements. *Alu* insertions have been reported in the literature as causing human genetic diseases. However, this is the first report of a pathogenic *CAPN3* gene mutation showing the contribution of an *Alu* insertion.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are indebted to the family for their invaluable cooperation and for providing the blood samples. This research was funded by the Tunisian Ministry of Higher Education and Scientific Research. We wish to thank Mr Jamil JAOUA, founder and former Head of the English Unit at the Sfax Faculty of Science, Tunisia, for proofreading this paper.

1 Richard, I., Broux, O., Allamand, V., Fougerousse, F., Chiannilkulchai, N., Bourg, N. *et al.* Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell* **81,** 27–40 (1995).
2 Garnham, P., Hanna, A., Chou, S., Low, E., Gourlay, K., Campbell, L. *et al.* Limb-girdle muscular dystrophy type 2A can result from accelerated autoproteolytic inactivation of calpain 3. *Biochemistry* **48,** 3457–3467 (2009).
3 Sorimachi, H., Imajoh-Ohmi, S., Emori, Y., Kawasaki, H., Ohno, S., Minami, Y. *et al.* Molecular cloning of a novel mammalian calcium-dependent protease distinct from both m- and mu-types. Specific expression of the mRNA in skeletal muscle. *J. Biol. Chem.* **264,** 20106–20111 (1989).
4 Suzuki, K., Hata, S., Kawabata, Y. & Sorimachi, H. Structure, activation, and biology of calpain. *Diabetes* **53,** S12–S18 (2004).
5 Sorimachi, H., Hata, S. & Ono, Y. Calpain chronicle—an enzyme family under multidisciplinary characterization. *Proc. Jpn. Acad. Ser. B* **87,** 287–327 (2011).
6 Sorimachi, H., Hata, S. & Ono, Y. Impact of genetic insights into calpain biology. *J. Biochem.* **150,** 23–37 (2011).
7 Groen, E. J., Charlton, R., Barresi, R., Anderson, L. V., Eagle, M., Hudson, J. *et al.* Analysis of the UK diagnostic strategy for limb girdle muscular dystrophy 2A. *Brain* **130,** 3237–3249 (2007).
8 Kramerova, I., Beckmann, J. S. & Spencer, M. J. Molecular and cellular basis of calpainopathy (limb girdle muscular dystrophy type 2A). *Biochim. Biophys. Acta* **1772,** 128–144 (2007).
9 Laval, S. H. & Bushby, K. M. Limb-girdle muscular dystrophies—from genetics to molecular pathology. *Neuropathol. Appl. Neurobiol.* **30,** 91–105 (2004).
10 Urtasun, M., Saenz, A., Roudaut, C., Poza, J J., Urtizberea, J A., Cobo, A M. *et al.* Limb-girdle muscular dystrophy in Guipuzcoa (Basque Country, Spain). *Brain* **121,** 1735–1747 (1998).
11 Fardeau, M., Eymard, B., Mignard, C., Tomé, F. M., Richard, I. & Beckmann, J. S. Chromosome 15-linked limb-girdle muscular dystrophy: clinical phenotypes in Reunion Island and French metropolitan communities. *Neuromuscul. Disord.* **6,** 447–453 (1996).
12 Pollitt, C., Anderson, L. V. B., Pogue, R., Davison, K., Pyle, A. & Bushby, K. M. The phenotype of calpainopathy: diagnosis based on a multidisciplinary approach. *Neuromuscul. Disord.* **11,** 287–296 (2001).
13 Fanin, M., Nascimbeni, A. C. H., Tasca, E. & Angelini, C. How to tackle the diagnosis of limb-girdle muscular dystrophy 2A. *Eur. J. Hum. Genet.* **17,** 598–603 (2009).
14 Beckmann, J. S. & Spencer, M. Calpain 3, the 'gatekeeper' of proper sarcomere assembly, turnover and maintenance. *Neuromuscul. Disord.* **18,** 913–921 (2008).
15 Nascimbeni, A. C. H., Fanin, M., Tasca, E. & Angelini, C. Transcriptional and translational effects of intronic CAPN3 gene mutations. *Hum. Mutat.* **31,** 1658–1669 (2010).
16 Blázquez, L., Azpitarte, M., Sáenz, A., Goicoechea, M., Otaegui, D., Ferrer, X. *et al.* Characterization of novel CAPN3 isoforms in white blood cells: an alternative approach for limb-girdle muscular dystrophy 2A diagnosis. *Neurogenetics* **9,** 173–182 (2008).
17 Krahn, M., Pécheux, C., Chapon, F., Béroud, C., Drouin-Garraud, V., Laforet, P. *et al.* Transcriptional explorations of CAPN3 identify novel splicing mutations, a large-sized genomic deletion and evidence for messenger RNA decay. *Clin. Genet.* **72,** 582–592 (2007).
18 Stehlíková, K., Zapletalová, E., Sedlácková, J., Hermanová, M., Vondrácek, P., Maríková, T. *et al.* Quantitative analysis of CAPN3 transcripts in LGMD2A patients: involvement of nonsense-mediated mRNA decay. *Neuromuscul. Disord.* **17,** 143–147 (2007).
19 Mattick, J. S. Introns: evolution and function. *Curr. Opin. Genet. Dev.* **4,** 823–831 (1994).
20 Christie, P. T., Harding, B., Nesbit, M. A., Whyte, M. P. & Thakker, R. V. X-linked hypophosphatemia attributable to pseudoexons of the PHEX gene. *J. Clin. Endocrinol. Metab.* **86,** 3840–3844 (2001).
21 Metherell, L. A., Akker, S. A., Munroe, P. B., Rose, S. J., Caulfield, M., Savage, M. O. *et al.* Pseudoexon activation as a novel mechanism for disease resulting in atypical growth-hormone insensitivity. *Am. J. Hum. Genet.* **69,** 641–646 (2001).
22 Chillon, M., Dork, T., Casals, T., Gimenez, J., Fonknechten, N., Will, K. *et al.* A novel donor splice site in intron 11 of the CFTR gene, created by mutation 1811 + 1.6 kbA–>G, produces a new exon: high frequency in Spanish cystic fibrosis chromosomes and association with severe phenotype. *Am. J. Hum. Genet.* **56,** 623–629 (1995).
23 Vervoort, R., Gitzelmann, R., Lissens, W. & Liebaers, I. A mutation (IVS8 + 0.6kbdelTC) creating a new donor splice site activates a cryptic exon in an *Alu*-element in intron 8 of the human beta-glucuronidase gene. *Hum. Genet.* **103,** 686–693 (1998).
24 Lewin, H. A. & Stewart-Haynes, J. A. A simple method for DNA extraction from leukocytes for use in PCR. *Biotechniques* **13,** 522–524 (1992).
25 Hadj Salem, I., Kamoun, F., Louhichi, N., Rouis, S., Mziou, M. & Fendri-Kriaa, N. Mutations in LAMA2 and CAPN3 genes associated with genetic and phenotypic heterogeneities within a single consanguineous family involving both congenital and progressive muscular dystrophies. *Biosci. Rep.* **31,** 125–135 (2011).
26 Schug, J. UNIT 2.6 using TESS to predict transcription factor binding sites in DNA sequence. *Curr. Protoc. Bioinform.* Chapter 2: Unit 2.6, doi:10.1002/0471250953.bi0206s21 (2008).
27 de Vooght, K. M., van Wijk, R. & van Solinge, W. W. Management of gene promoter mutations in molecular diagnostics. *Clin. Chem.* **55,** 698–708 (2009).
28 Reese, M. G., Eeckman, F. H., Kulp, D. & Haussler, D. Improved splice site detection in Genie. *J. Comput. Biol.* **4,** 311–323 (1997).
29 Desmet, F. O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M. & Béroud, C. Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37,** e67 (2009).
30 Churbanov, A., Vořechovský;, I. & Hicks, C. A method of predicting changes in human gene splicing induced by genetic variants in context of cis-acting elements. *BMC. Bioinformatics* doi:10.1186/1471-2105-11-2 **11,** 22 (2010).
31 Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q. & Krainer, A. R. ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* **31,** 3568–3571 (2003).
32 De Paula, F., Vainzof, M., Passos-Bueno, M. R., de Cássia, M. P. R., Matioli, S. R., Anderson, V. B. L. *et al.* Clinical variability in calpainopathy: what makes the difference? *Eur. J. Hum. Genet.* **10,** 825–832 (2002).
33 Chae, J., Minami, N., Jin, Y., Nakagawa, M., Murayama, K., Igarashi, F. *et al.* Calpain 3 gene mutations: genetic and clinico-pathologic findings in limb-girdle muscular dystrophy. *Neuromuscul. Disord.* **11,** 547–555 (2001).
34 Rousseau, P., Le Discorde, M., Mouillot, G., Marcou, C., Carosella, E. D. & Moreau, P. The 14 bp deletion-insertion polymorphism in the 3′UT region of the HLA-G gene influences HLA-G mRNA stability. *Hum. Immunol.* **64,** 1005–1010 (2003).
35 Di Liegro, C. M., Bellafiore, M., Izquierdo, J. M., Rantanen, A. & Cuezva, J. M. 3′-untranslated regions of oxidative phosphorylation mRNAs function *in vivo* as enhancers of translation. *Biochem. J.* **352,** 109–115 (2000).
36 Piluso, G., Politano, L., Aurino, S., Fanin, M., Ricci, E., Ventriglia, V. M. *et al.* Extensive scanning of the calpain-3 gene broadens the spectrum of LGMD2A phenotypes. *J. Med. Genet.* **42,** 686–693 (2005).
37 De Tullio, R., Stifanese, R., Salamino, F., Pontremoli, S. & Melloni, E. Characterization of a new p94-like calpain form in human lymphocytes. *Biochem. J.* **375,** 689–696 (2003).
38 Kawabata, Y., Hata, S., Ono, Y., Ito, Y., Suzuki, K., Abe, K. *et al.* Newly identified exons encoding novel variants of p94/calpain 3 are expressed ubiquitously and overlap the a-glucosidase C gene. *FEBS Lett.* **555,** 623–630 (2003).
39 Aissi-Ben Moussa, S., Moussa, A., Lovecchio, T., Kourda, N., Najjar, T., Ben Jilani, S. *et al.* Identification and characterization of a novel MLH1 genomic rearrangement as the cause of HNPCC in a Tunisian family: evidence for a homologous Alu-mediated recombination. *Fam. Cancer* **8,** 119–126 (2009).
40 Pastor, T., Talotti, G., Lewandowska, M. A. & Pagani, F. An *Alu*-derived intronic splicing enhancer facilitates intronic processing and modulates aberrant splicing in ATM. *Nucleic Acids Res.* **37,** 7258–7267 (2009).
41 Lev-Maor, G., Ram, O., Kim, E., Sela, N., Goren, A., Levanon, E. Y. *et al.* Intronic Alus influence alternative splicing. *PLoS Genet.* **4,** e1000204 (2008).
42 Hasler, J. & Strub, K. Alu elements as regulators of gene expression. *Nucleic Acids Res.* **34,** 5491–5497 (2006).
43 Mahadevan, S., Foitzik, A., Surh, C. & Korneluk, G. Characterization and polymerase chain reaction (PCR) detection of an Alu deletion polymorphism in total linkage disequilibrium with myotonic dystrophy. *Genomics* **15,** 446–448 (1993).
44 Varon, R., Gooding, R., Steglich, C., Marns, L., Tang, H., Angelicheva, D. *et al.* Partial deficiency of the C-terminal-domain phosphatase of RNA polymerase II is associated with congenital cataracts facial dysmorphism neuropathy syndrome. *Nat. Genet.* **35,** 185–189 (2003).

45 Zhang, Y., Dipple, K. M., Vilain, E., Huang, B. L., Finlayson, G., Therrell, B. L. *et al.* AluY insertion (IVS4-52ins316alu) in the glycerol kinase gene from an individual with benign glycerol kinase deficiency. *Hum. Mutat.* **15,** 316–323 (2000).
46 Oldridge, M., Zackai, E. H., McDonald-McGinn, D. M., Iseki, S., Morriss-Kay, G. M., Twigg, S. R. *et al. De novo* alu-element insertions in FGFR2 identify a distinct pathological basis for Apert syndrome. *Am. J. Hum. Genet.* **64,** 446–461 (1999).
47 Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, P. E., Glover, T. W. & Collins, F. S. A *de novo* Alu insertion results in neurofibromatosis type 1. *Nature* **353,** 864–866 (1991).
48 Ganguly, A., Dunbar, T., Chen, P., Godmilow, L. & Ganguly, T. Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. *Hum. Genet.* **113,** 348–352 (2003).
49 Tighe, P. J., Stevens, S. E., Dempsey, S., Le Deist, F., Rieux-Laucat, F. & Edgar, J. D. Inactivation of the Fas gene by Alu insertion: retrotransposition in an intron causing splicing variation and autoimmune lymphoproliferative syndrome. *Genes Immun.* **3,** S66–S70 (2002).
50 Sorek, R., Ast, G. & Graur, D. Alu-containing exons are alternatively spliced. *Genome Res.* **12,** 1060–1067 (2002).
51 Makalowski, W., Mitchell, G. A. & Labuda, D. Alu sequences in the coding regions of mRNA: a source of protein variability. *Trends Genet.* **10,** 188–193 (1994).
52 Nekrutenko, A. & Li, W. H. Transposable elements are found in a large number of human protein-coding genes. *Trends Genet.* **17,** 619–621 (2001).
53 Massari, M. E. & Murre, C. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell Biol.* **20,** 429–440 (2000).
54 Fisher, A. L., Ohsako, S. & Caudy, M. The WRPW motif of the hairy-related basic helix–loop–helix repressor proteins acts as a 4-amino-acid transcription repression and protein–protein interaction domain. *Mol. Cell Biol.* **16,** 2670–2677 (1996).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)