

*MUTATION IN BRIEF*

# Homeologous Recombination Between AluSx-Sequences as a Cause of Hemophilia

Liliana C. Rossetti<sup>1\*</sup>, Anne Goodeve<sup>2</sup>, Irene B. Larripa<sup>1</sup>, and Carlos D. De Brasi<sup>1</sup>

<sup>1</sup> Instituto de Investigaciones Hematológicas Mariano R. Castex, Academia Nacional de Medicina de Buenos Aires, Argentina; <sup>2</sup> Division of Genomic Medicine, Royal Hallamshire Hospital, University of Sheffield, United Kingdom

\*Correspondence to: Liliana C. Rossetti, Departamento de Genética, Instituto de Investigaciones Hematológicas Mariano R. Castex, Academia Nacional de Medicina, Pacheco de Melo 3081, Buenos Aires 1425, Argentina; E-mail: rossetti@hematologia.anm.edu.ar

Grant sponsors: René Barón Foundation; Alberto J. Roemmers Foundation; National Research Council (CONICET); ANPCyT Argentina; The Royal Hallamshire Hospital, UK; World Federation of Hemophilia.

*Communicated by Steve Sommer*

**Although large deletions from the coagulation factor VIII gene, F8, are responsible for 5% of severe hemophilia A (seHA), few of them have been fully characterised. A detailed description of a large partial deletion of the F8 caused by unequal recombination between homeologous AluSx-derived sequences is presented. The proband, a case of isolated hemophilia A with a high inhibitor titre (5700 BU), showed a consistent absence of PCR-amplification of exons 4 to 10, EX4\_EX10del. Two approaches were used to narrow down the deletion breakpoints: a direct physical analysis based on PCR (that additionally permits carrier detection in the family); and, under the hypothesis that the mutation resulted from homologous recombination, sequence alignments of F8 intron 3 and 10. Both approaches indicate an unequal crossing over (CO) between two Alu-related sequences. Both elements involved were derived from the AluSx-subfamily consensus and demonstrate 86% sequence identity (with only single-base mismatches), with three gaps (of 2, 3 and 14-bases) and two main tracts of perfectly homologous sequence (28 and 24-bp). The short stretch of intron 10 embedded into intron 3 sequence, linked to the CO, represents a typical hallmark of homologous recombination (double-strand break repair model). A detailed description of EX4\_EX10del mutation is c.[338+3485delins1687+2223\_1687+2225; 338+3551\_1687+2291 del]. The common involvement of unequal homologous recombination mediated by repetitive elements allowed us to suggest that our experimental design (based on intron sequence alignments) may be successfully applied to rearrangements involved in other X-linked inherited diseases. Like other Alu-rich genes throughout the human genome, Alu-mediated homologous recombination in F8 may be an important cause of hemophilia by promoting large DNA deletions. © 2004 Wiley-Liss, Inc.**

KEY WORDS: F8; hemophilia A; HEMA; mutation; deletion; Alu repeats.

## INTRODUCTION

Recombination between homologous DNA sequences plays an important role in generating genetic diversity in

Received 25 May 2004; accepted revised manuscript 24 August 2004

all organisms. Similar to homologous recombination, where both sequences involved are almost identical, homeologous recombination involves similar but imperfectly matched sequences (Yang & Waldman, 1997). Although the limit for the above classification is arbitrary, it has been shown that disruption of sequence identity greatly reduces the chance of recombination between homologs (Waldman & Liskay, 1987). Short interspersed elements, such as Alu-elements, have propagated via retrotransposition to more than one million copies in the human genome, thus providing numerous possibilities for unequal recombination (Deininger & Batzer, 1999). Old Alu-subfamilies such as AluJ, AluSx and AluSq have been shown to harbour a large number of random mutations, which prevent their retrotransposition and also reduce their chance of unequal homologous recombination leading to human disease (Batzer et al., 1996).

Hemophilia A (MIM# 306700) is an X-linked recessive disorder, which affects one in 10,000 males world-wide and is caused by mutations affecting the coagulation factor VIII (FVIII) gene (F8). The disease is classified as severe (seHA) when the coagulant activity of plasma FVIII is lower than 1 IU/dl. F8 is large and complex (186 kb and 26 exons) and is situated in the distal Mb of Xq (Gitschier et al., 1984). The first recurrent F8 mutation to be recognised was the intron 22 inversion (Inv22) (Lakich et al., 1993; Naylor et al., 1993), which accounts for approximately 42% of seHA cases (Antonarakis et al., 1995; De Brasi et al., 2000). Inv22 is mediated by intrachromosomal homologous recombination between a region of 9.5 kb within F8 IVS22 (*int22h-1*) and one of two inversely oriented copies of this sequence (*int22h-2* or *-3*) (Naylor et al., 1995). Recently, Bagnall et al. (2002) described another recurrent inversion involving F8, affecting IVS1 (Inv1), that is responsible for up to 5% of seHA cases and results from homologous recombination between inversely oriented sequences of 1.2 kb (*int1h*). Large deletions in the F8 account for approximately 5% of seHAs (HAMSTeRS international hemophilia A mutation database <http://europium.csc.mrc.ac.uk>, Kemball-Cook et al., 1998). Genetic counselling in families with large F8 deletions can become a priority as a result of the increased clinical severity of the proband as a result of the high probability of developing an anti-FVIII inhibitory antibody (inhibitor). Carrier analysis in deletion families is hampered by the presence of a masking normal allele in carrier females. In this context, rapid analysis of a deletion-specific detection system may be extremely valuable. F8 is an Alu-rich gene with at least 49 Alu-core related sequences. It has been speculated that these sequences may provide many possibilities for unequal recombination (Vidal et al., 2002; Sommer et al., 2002). Although several papers have reported large partial deletions in F8 (Yousoufian et al., 1987; Bardoni et al., 1988; Millar et al., 1990; Krepelova et al., 1992; Figueredo et al., 1994; Salviato et al., 2002), few of them have presented a complete characterisation of the sequence breakpoints and addressed the mechanism involved (Woods-Samuels et al., 1991; Van de Water et al., 1998; Shibata et al., 2000, Vidal et al., 2002).

Although the origin of large deletions in F8 was formerly shown to be associated with non-homologous recombination (Woods-Samuels et al., 1991), it is currently believed to involve alternative molecular mechanisms. Recently, the involvement of a 16 bp-subset of the recombinogenic Alu-core sequence (Rüdiger et al., 1995) in a F8 exon 25 deletion (Vidal et al., 2002) prompted Sommer et al. (2002) to raise a central question of whether Alu-mediated recombination is a common mechanism of mutation in the FVIII gene.

In order to address this question and to gain further insight into the molecular mechanism of Alu-mediated rearrangements in the human genome, we studied a large partial deletion of F8 associated seHA. Herein, we provide evidence that this deletion is caused by unequal homeologous recombination between two full-length copies of AluSx situated in introns 3 and 10 of F8.

## MATERIALS AND METHODS

### Case and samples

The proband was a patient with sporadic severe HA (FVIII:C < 1 IU/ml) who presented with a high titre inhibitory antibody against therapeutic FVIII (titre: first test 14,4 BU; maximum 5700 BU). The proband, his mother and 2 unrelated healthy women were studied to characterize the mutation, for carrier detection, and as control individuals. Genomic DNA was obtained from peripheral blood leukocytes either by phenol-chloroform (Sambrook et al., 1989) or salting-out extraction (Lahiri & Nuremberg, 1990) and ethanol precipitation. DNA quality and concentration were estimated by agarose gel electrophoresis.

**DNA sequences**

Xq28 contig, GenBank accession AL645722.2, AluSx, AluRep accession U14574. HEMOA: OMIM ID 306700; International hemophilia A mutation database: HAMSTeRS (<http://europium.csc.mrc.ac.uk>).

**Intron 22 and intron 1 inversion analysis**

The intron 22 inversion was investigated by Southern blot analysis as described Lakich et al. (1993). The intron 1 inversion was studied by the method described by Bagnall et al. (2002) with minor modifications.

**Long distance and standard PCR amplification**

Long-distance PCR amplifications (LD-PCR) were performed on 500 ng of genomic DNA, in a total volume of 25µl, using 1U of *Taq/Pwo* DNA polymerase mixture (Expand long template DNA polymerase Mix, Roche) as previously described (De Brasi et al., 1999), with minor modifications using primers 3a and 11b for the deletion; and 3a and 4b for the wt-allele (Table 1, Fig. 1a). Deletion-specific and wild type intron-specific short range PCR amplification (S-PCR) were achieved on 300 ng of DNA, in a volume of 25 µl, using standard reagents and thermo-cycling, with primers i3f and i10r; and with i3f and 4b; respectively (Table 1, Fig. 1a). For point mutations screening, all F8 coding sequences were PCR-amplified from genomic DNA samples as described by Williams et al. (1998).

**Table 1. Primers Used for PCR Amplification**

Code	5' – Sequence – 3'	GenBank AL645722.2 Nucleotides	PCR Assay
3a	TGCTTCTCCACTGTGACCT *	44962-44980	LD-PCR
4b	GATTCAGTTGTTTGTACTTCTC *	49258-49279	LD-PCR S-PCR
11b	ACTGACCTATATTGCAAACCA *	85330-85350	LD-PCR
i3f	TTTGTGGCTGCTGCTTGTAGT **	48556-48577	S-PCR
i10r	AAAATCTCCCCAGAACTCTGTCA **	83669-83691	S-PCR

\* Williams et al. (1998). \*\* This study.

**Pre-screening and screening for small deletions of factor VIII gene**

F8 gene mutation screening of all 26 exons, exon-intron boundaries and splicing sequences (260- 547bp), was performed using conformation sensitive gel electrophoresis (CSGE) as described Williams et al. (1998). Briefly, products from the clinical sample and a non-mutated control were mixed to form heteroduplexes, electrophoresed in mildly denaturing polyacrylamide gels, ethidium bromide stained and analyzed by UV transillumination.

**Restriction endonuclease analysis and electrophoresis of PCR products**

PCR products (5-10µl) were digested using 5U of restriction enzyme in 15µl total volume under conditions recommended by the manufacturer (Promega, Gibco BRL, New England Biolabs). LD-PCR products and their restriction digestion products were analyzed using 0.8-1% agarose gel electrophoresis while short-range PCR products and their restriction digests, were analyzed by 1.5-3% agarose gel electrophoresis. Gel images were documented using GelPro analyzer and the molecular weight of DNA segments was estimated by GelPro 3.2 analysis software (Syrex, Argentina).

**Bioinformatics**

DNA sequence analysis, restriction mapping, exon-intron boundaries and primer annotations were performed using EditSeq and MapDraw software (LaserGene, DNA Star). Intron alignments were achieved by the Clustal

algorithm using default penalty parameters unless stated (MegAlign, Lasergene, DNA Star). Restriction maps to characterize the large deletion were achieved using an updated and F8 exon-annotated genomic sequence of Xq28, GenBank accession AL645722.2. The NCBI web page (<http://www.ncbi.edu/blast/>) was used to access BLAST (Basic Local Alignment Search Tool) algorithm (Atschul et al., 1990). AluSx subfamily sequence consensus were obtained from AluRep, accession U14574 (Batzer et al., 1996).

#### DNA sequencing

PCR products were prepared for DNA sequencing by purification using Concert Gel Extraction System (Gibco BRL). Sequencing primers were those used for PCR amplifications (Table 1). Manual DNA sequencing was performed using *fmol* DNA Cycle Sequencing System (Promega) under conditions recommended by the manufacturer. Each DNA sequence was determined using both forward and reverse primers.

## RESULTS

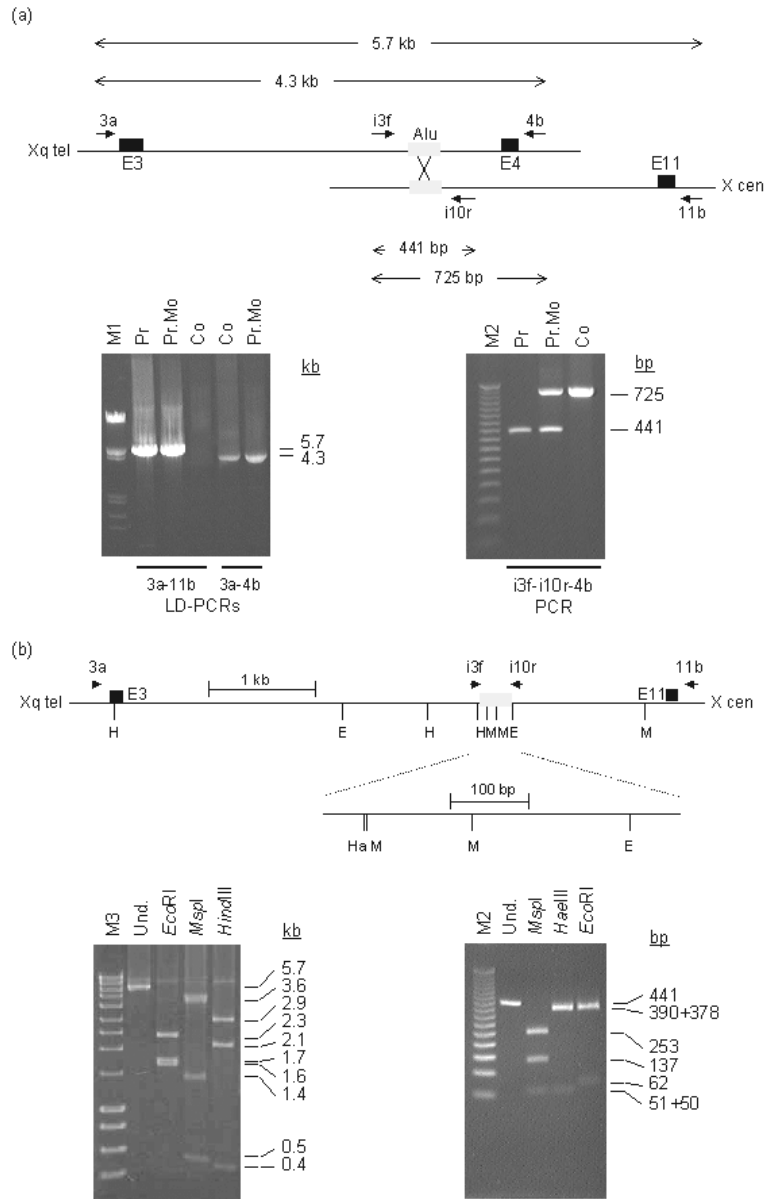
#### Identification of a large deletion

The seHA-affected patient was shown not to have the Inv22 or Inv1. The proband's DNA sample was successfully amplified for exons 1 to 3 of F8, and exons 11 to 26. Exons 4-10 consistently failed to amplify. The amplimers obtained were screened for small mutations (small deletions, insertions and single base changes), by CSGE. The only change identified was in nucleotide 3864A>C, codon 1269 (Ser>Ser), previously reported on HAMsTERS as a common exon 14 DNA polymorphism. The results suggested a large partial deletion of F8 spanning exons 4 to 10.

#### Characterization of the sequence breakpoints at the genomic level

In order to isolate the genomic breakpoints associated with the EX4\_EX10del, a LD-PCR was performed using primers 3a and 11b, (Table 1) specific to the exons flanking the deletion (Fig. 1a). Interestingly, a deletion-specific product of 5.7 kb was produced in samples from the patient and his mother but not in either control individual's DNA (Fig. 1a). At this point a hypothesis of recombination between homologous segments of introns 3 and 10 prompted us to perform a sequence alignment and to formulate a prediction of the breakpoint sequences. This indicated that two Alu-related sequences may be involved in homologous recombination between introns 3 and 10 of the F8, and predicted an amplimer of 5694 bp using primers 3a and 11b. Analysis of intron 3 and intron 10 parental Alu homologs using BLASTN on the AluRep database revealed that both Alu-elements were derived from the AluSx subfamily. Due to sequence divergence, these parental elements showed a moderate number of random mutations compared with AluSx consensus (Fig. 2).

Restriction endonuclease analysis of the 5.7-kb amplicon was compatible with the predicted Alu-mediated CO event within the DNA segments indicated (Fig. 1b). These results encouraged us to perform a short-range PCR specific for the deletion (441 bp) accompanied with another S-PCR to detect only the non-deleted allele. This new system based on S-PCR succeeded in detecting the EX4\_EX10del and was used to simplify further studies in the family to confirm the carrier status of the proband's mother (Fig. 1a), and to characterize the deletion breakpoints at the DNA sequence level. Restriction analysis of this 441-bp amplimer was also compatible with the predictions made by bioinformatics (Fig. 1b). The DNA sequence through the breakpoint of EX4\_EX10del was determined using the 441-bp PCR product (Fig. 2).



**Figure 1.** Isolation and analysis of EX4\_EX10del. (a) Above. Scheme of the deletion showing AluSx (gray boxes) mediated recombination. Primer target sites and PCR-product sizes are indicated. Below. Long distance PCR (left panel) for genotyping EX4\_EX10del (primers 3a and 11b) and the normal allele (primers 3a and 4b) samples from the proband (Pr), the proband's mother (Pr.Mo) and a normal individual (Co); and in the right panel, the same analysis by multiplex standard PCR (S-PCR) (primers i3f, i10r and 4b). These two analysis show that the proband and his mother carry the specific deletion. M1 ( $\lambda$ -EcoRI-HindIII), M2 (50-bp ladder). (b) Restriction maps from the proband to analyze EX4\_EX10del specific amplimers. H (HindIII), E (EcoRI), M (MspI), Ha (HaeIII). This map was drawn using genomic sequence data (GenBank accession AL645722.2) assuming unequal crossing over between Alu-elements within IVS3 and IVS10 of F8. Below. Electrophoretic analysis of DNA fragments obtained from the proband's 5.7-kb LD-PCR product (primers 3a and 11b) (left) and the 441-bp S-PCR product (primers i3f and i10r) (right). M3 (1-kb hyperladder), Und. (undigested PCR-product). The restriction maps that were obtained perfectly coincided with those that were predicted assuming homologous Alu mediated recombination.

```

1  GGC GGGCGGA  GGC GGGCGC  GGTGGTCC  G C C T G T A A T C  C C A G C A C T T T  G G G A G G A ] - - - - - - - A G A  T C A C C T G A G G  A l u S x
+3411 G C A G A A A G T A G G C C G G G T G T G G T G G T C G C G C C T G T A A T C C C A G C A C T T T G G G A G G C T G A G G C G G G T G G A T C A C - - A A G G  I V S 3
+2149 A A A A T T T T C C C G C T G G G T G C G G T G G T C A C G C C T G T A A T C C C A G C A C T T T G G G A G G C C G A G G C G G G C A G A T C A C T T G A G G  I V S 10
1  G C A G A A A G T A  G G C C G G G T G T  G G T G G T C G C  G C C T G T A A T C  C C A G C A C T T T  G G G A G G C T G A  G G C G G G T G G A  T C A C T T G A G G  E X 4 _ E X 10 d e l

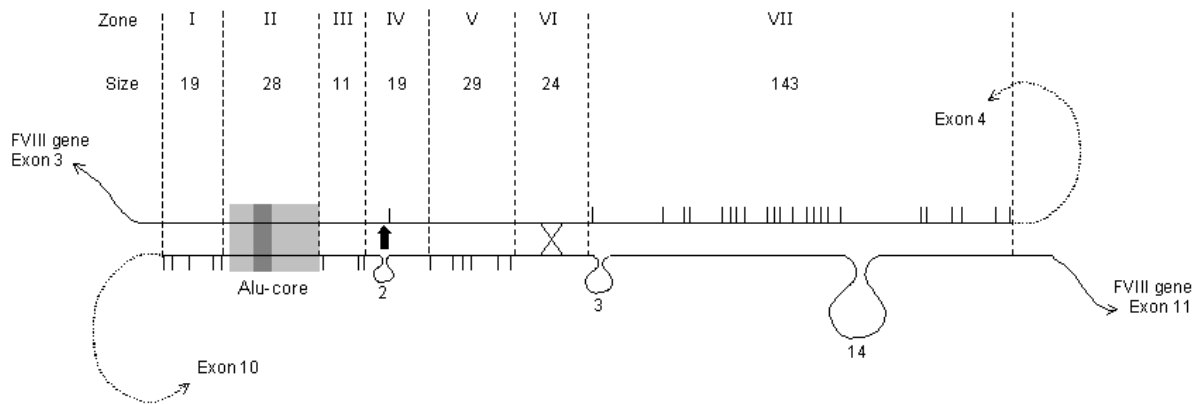
71  T C A G G A G T T C  G A G A C C A G C C  T G G C C A A C A T  G G T G A A A C C C  C G T C T C T A C T  A A A A A T A C A A  A A A T - - T A G C  C G G C G T G G T  A l u S x
+3490 T C A G G A G T T T  A A G A C C A T C C  T G A C C A A C A T  G A T G A A A C C C  C A T C T C T A C T  A A A A A T A C A A  A A T - - - T A G C  C G G G C A T G G T  I V S 3
+2230 T C A G G A G T T C  A A G A C C A G C T  T G G C C A A C A T  G G T G A A A A C C  C A T C T C T A C T  A A A A A T A C A A  A A A A A A T A G C  C G G G C A T T G T  I V S 10
81  T C A G G A G T T T  A A G A C C A T C C  T G A C C A A C A T  G A T G A A A C C C  C A T C T C T A C T  A A A A A T A C A A  A A A A A A T A G C  C G G G C A T T G T  E X 4 _ E X 10 d e l

149  G G C G C G C G C C  T G T A A T C C C A  G C T A C T C G G G  A G G C T G A G G C  A G G A G A A T C G  C T T G A A C C C G  G G A G G C G G A G  G T T G C A G T G A  A l u S x
+3567 G G C A C G T G C C  T G T A A T C C C A  G C T G C T G G G G  A G T C T G A G G C  A G G A A A A T C A  C T T G A A C C T G  G G A G G C G G A C  G T T G C T G T G A  I V S 3
+2310 G G C G T G T G C C  T G T A A T C C C A  G C T A C T C A G G  A G G C T G A G G C  A G G A G A A T T G  C T T G A A A C T G  G G A G G T G A A G  G T T G C A G T G A  I V S 10
161  G G C G T G T G C C  T G T A A T C C C A  G C T A C T C A G G  A G G C T G A G G C  A G G A G A A T T G  C T T G A A A C T G  G G A G G T G A A G  G T T G C A G T G A  E X 4 _ E X 10 d e l

229  G C C G A G A T C G  C G C C A C T G C A  C T C C A G C C T G  G G C G A C A G A G  C G A G A C T C C G  T C T C A A A A A A  A A 290 A l u S x
+3647 G C C G A G A T C - - - - - - - - - - - - - C A G C C T G  G G C A A C A G A G  T G A G A C T C C G  T C T C A A A A C A  A C 3694 I V S 3
+2390 G C C A A G A T C A  C C C C A C T G C A  C T C C A G C C T G  G G C A A C A G A G  C A A G A C T G C A  T C T C A A A A A A  A A 2451 I V S 10
241  G C C A A G A T C A  C C C C A C T G C A  C T C C A G C C T G  G G C A A C A G A G  C A A G A C T G C A  T C T C A A A A A A  A A 302 E X 4 _ E X 10 d e l

```

**Figure 2.** Sequence alignment of the full-length AluSx consensus (GenBank accession U14574) (AluSx) (Batzer et al., 1996), the parental IVS3 position c.338+3411 to c.338+3694 (GenBank accession AL645722.2: 48595-48878), the parental IVS10 position c.1687+2149 to c.1687+2451 (GenBank accession AL645722.2: 83287-83589) and the chimeric DNA sequence from EX4\_EX10del. Sequence gaps are indicated by [-], bold type indicates base disagreement between IVS3 and IVS10; underlined letters represent base disagreements with the AluSx consensus. The open box indicates Alu-core sequence and the shaded box, the Chi-like motif.



**Figure 3.** Hypothetical scheme for mechanism of recombination that generated the EX4\_EX10del mutation. The best alignment of F8 IVS3 and IVS10 involves two almost full-length AluSx elements. This alignment starts at nucleotide IVS3 +3421 and IVS10 +2159, and ends at IVS3 +3694 and IVS10 +2451 (figure 2). Solid lines are sequences represented in EX4\_EX10del, while dashed lines are not; short vertical lines indicate mismatched bases between either the IVS3 or IVS10 Alu-element, and the observed nucleotide sequence of EX4\_EX10del; loops are sequence gaps. The tract of DNA sequence homology (292 bp including gaps) is arbitrarily divided into seven zones (I, II, ..., VII). Segment sizes in nucleotides is indicated. Under the most likely hypothesis, the cross over (CO) [X] occurred in zone VI associated with the conversion of a short sequence within zone IV to the corresponding sequences from IVS10, indicated by the upward arrow (at least 3 bases, IVS3 nt. +3485 "A" was replaced by "TTG" of IVS10 nt. +2223 to +2225 on figure 2). However, other molecular mechanisms cannot be ruled out. For example the CO could have been located at the most 5' six nucleotides of zone IV (from the last mismatch of zone III up to the 2-bp gap, IVS3 nt +3479 to +3484, figure 2) associated with the conversion (a patch of IVS3 sequence as information donor) of more than 29 bases spanning the entire zone V. Alternative, more complex molecular outcomes (for example a CO associated with two separated gene conversion-like events) reconcile less well with the current theoretical background (Szostak et al., 1983).

A formal and precise description of the deletion associated sequences has been assessed according current recommendations (Dunnen & Antonarakis, 2000), as c.[338+3485delins1687+2223\_1687+2225; 338+3551\_1687+2291del]. These data provide conclusive evidence that the seHA-causing EX4\_EX10del large deletion originated by recombination between two AluSx-derived sequences.

#### Analysis of sequence features associated with the EX4\_EX10del

Alignment of the parental intron 3 and intron 10 Alu-elements involved in the origin of the deletion showed the presence of 38 single-base mismatches, two tracts of sequence identity (one of 28 bases spanning the Alu-core and the Chi-like pentanucleotide and the other of 24 bp) and three gap-loops each of 2, 3 and 14 nucleotides, respectively; through the 292 bp of sequence homology (including gaps) (Figs. 2 and 3). The overall figure for sequence identity between the homeologous parental tracts involved in EX4\_EX10del was 81% and reached 86% when gaps were excluded and only single-base mismatches were considered (Fig. 3, sequence details in Fig. 2). This degree of nucleotide sequence identity classified the sequences as homeologous (Nicholson et al., 2000).

In order to describe possible outcomes of the event, the homeologous tracts were arbitrary divided into seven zones (I, II, III, IV, V, VI and VII) (Fig. 3). Under the most likely hypothesis, the CO occurred within zone VI associated with the ectopic presence of an area of IVS10 embedded in IVS3 sequence (zone IV, Fig. 3). Although no direct evidence of this hypothesis can be obtained, in order to rule out the possibility of errors in the published sequence of F8 IVS3 in zone IV (GenBank accession AL645722.2: 48661-48682) that could mimic IVS10 zone IV sequence; four X-chromosomes from healthy individuals were sequenced. Only the published DNA sequence was obtained.

## DISCUSSION

This study presents the second report of a partial genomic deletion causing hemophilia in which Alu-sequences were involved (Vidal et al., 2002) and the first one where full-length Alu-elements were identified. The event reported here is particularly interesting as it involved two divergent copies of the AluSx subfamily and where sequence differences between the repeats allowed us to further explore some aspects of homeologous recombination in human meiosis.

Alu repeats represents over 10% of the mass of the human genome (Lander et al., 2001), but repeat density varies greatly. Evidence has been presented in yeast that retrotransposable elements may aid in healing chromosomal breaks via homologous recombination (Moore & Haber, 1996). In addition to the potential impact of Alu element insertions, it has been shown that the abundance of Alu-elements provides several opportunities for unequal homologous recombination to cause disease (Deininger & Batzter, 1999). Moreover, Rudiger et al. (1995) have shown that a 26 bp region of Alu (the Alu-core) is associated with both homologous and non-homologous recombination breakpoints. Since then, a number of studies have reported the involvement of Alu-Alu recombination in causing cancer as well as in germ-line disease including  $\alpha$ -thalassaemia ( $\alpha$ -globin) (Harteveld et al., 1997), hypercholesterolemia (LDLR), Ehlers-Danlos syndrome (Lys hydroxylase), thrombophilia (antithrombin); (reviewed in Deininger & Batzter, 1999), Lesch-Nyhan syndrome (HPRT) (Tvrdik et al., 1998), hereditary non polyposis colorectal cancer (mismatch repair MLH1) (Viel et al., 2002) and Ewing's sarcoma gene (EWS) (Zucman-Rossi et al., 1997) among other Alu-rich genes, in which Alu-Alu recombination is frequently involved in causing disease. To our knowledge, these studies have never focused on the extent and type of sequence identity among the Alu-homologs involved. Most of these studies found a relatively short tract of sequence identity (often reduced to the Alu-core or a subsequence of it) to which the single CO could be narrowed-down; or alternatively, full-length Alu-copies with a high degree of homology (> 90-95%). Under the hypothesis that the full-length Alu-partners play the leading role in promoting the recombinogenic event; with 86% of sequence identity through 292 bp and 3 gaps of different length, our patient demonstrates how quite divergent copies of Alu can still recombine using the homologous recombination mechanism (Fig. 3). Seventeen of 38 of the single-base mismatches between the Alu-sequences from IVS3 and IVS10 were related to CpG dinucleotides, thus reinforcing the observation that this is the main mechanism of sequence divergence and helps to prevent aberrant unequal rearrangements between duplicons (Krickler et al., 1992). However, the case reported here demonstrates that Alu sequences divergences (average homology among old Alu-subfamilies in the human genome is 81% overall), is unable to completely abolish events of unequal homologous recombination. On the other hand the possibility that the CO event was mediated entirely by the sequence identity provided by the Alu-core (embedded in zone II, Fig. 3) as has been observed by others (Vidal et al., 2002) cannot be dismissed.

Lopes et al. (1999) provided robust evidence associating the homologous DNA exchange in humans with the yeast double-strand break repair (DSBR) model (Szostak et al., 1983). This model provides the theoretical background to explain a hallmark associated with this specific EX4\_EX10del cross over; a short stretch of sequence conversion probably confined to zone IV (Fig. 3). In the context of meiotic homologous recombination, the association of CO with a tract of gene conversion is not uncommon (Borts & Haber, 1989). Although less likely, alternative predictions can be made to explain the deletion's origin (Fig. 3). In the DSBR model for homologous recombination, the CO may occur by resolution of two Holliday junctions, and the gene conversion may arise either obligated within the region of the gap produced by extension of the double-strand break, or facultative, by the repair of DNA heteroduplex intermediates formed at either side of the gap (Sztosak et al., 1983). This model provides a likely explanation for IVS10 specific sequence embedded into a tract of IVS3 sequence. Several studies have indicated that Alu-elements, as well as other mobile elements, undergo a large amount of gene conversion (Deininger & Batzer, 2002). Although associated with unequal CO, the stretch of sequence from one parental Alu-element between two DNA segments from the other parental Alu observed in the EX4\_EX10del; may be added as further indirect evidence of the phenomenon of Alu gene conversion already mentioned.

At the clinical level, the deletion reported here causes severe HA associated with the most serious complication for HA therapy, the development of inhibitory antibodies directed against therapeutic FVIII. The analysis of the HA mutation database permits the assignment of the highest risk of inhibitor development to patients with partial deletions involving more than one F8 exon (67%) (Goodeve et al., 2003). As predicted, our proband has developed an inhibitor.

The common involvement of unequal recombination among homologs as a cause of disease points to the utility of performing DNA sequence-alignments as these may provide accurate predictions for the bench experiments. Inspection of flanking intron sequence in other patients with partial F8 deletions may also suggest involvement of Alu sequences in the mutations. Design of deletion specific primers based on such prediction may enable deletion specific amplification analysis, which is particularly important for carrier diagnosis in female relatives of the patient. Our data support the hypothesis that unequal recombination of homologous Alu-related sequences in F8 may be an important cause of HA by promoting large partial deletions. Large deletions involved in hemophilia may therefore originate by an Alu-mediated homologous mechanism (Vidal et al., 2002; this study) as well as by non-homologous recombination (Wood-Samuels, 1990; Shibata et al., 2000).

#### **ACKNOWLEDGMENTS**

The authors thanks the haematologists Miguel Candela, Raúl Pérez Bianco and Miguel de Tezanos Pinto; as well as María Teresa Cuello, Patricia Gargallo and María de los Angeles Curto for their valuable help in different phases of the work.

#### **REFERENCES**

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403-410.
- Antonarakis SE and a consortium of more than 50 international authors. 1995. FVIII gene inversions in severe haemophilia A: results of an international consortium study. *Blood* 86: 2206-12.
- Bagnall RD, Waseem N, Green PM, Giannelli F. 2002. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe Hemophilia A. *Blood* 99: 168-74.
- Bardoni B, Sampietro M, Romano M, Crapanzano M, Mannucci PM, Camerino G. 1988. Characterization of a partial deletion of the factor VIII gene in a haemophiliac with inhibitor. *Hum Genet* 79: 86-88.
- Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM, Schmid CW, Zietkiewicz E, Zuckerkandl E. 1996. Standardized nomenclature for Alu repeats. *J Mol Evol* 42: 3-6.
- Borts RH, Haber JE. 1989. Length and distribution of meiotic gene conversion tracts and crossovers in *Saccharomyces cerevisiae*. *Genetics* 123: 69-80.
- De Brasi CD, Bowen DJ, Collins PW, Larripa IB. 1999. Specific analysis of the intron 22 XbaI polymorphism of the human factor VIII gene using long distance PCR. *British J Haematol* 107: 566-568.



- De Brasi CD, Candela M, Cermelj M, Larripa IB, Pérez Bianco R, de Tezanos Pinto M. 2000. Intron 22 factor VIII gene inversions in Argentine patients with severe haemophilia A. *Haemophilia* 6:21-2.
- Deininger PL, Batzer MA. 1999. Alu repeats and human disease. *Mol Genet Metab* 67: 183-193.
- Deininger PL, Batzer MA. 2002. Mammalian retroelements. *Genome Res* 12: 1455-1465.
- Dunnen JT den, Antonarakis SE. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion. *Hum Mutat* 15: 7-12.
- Figueiredo MS, Tavella MH, Simoes BP. 1994. Large DNA inversions, deletions and TaqI site mutations in severe haemophilia A. *Hum Genet* 94: 473-478.
- Gitschier J, Wood WI, Goralka TM, Wion KL, Chen EY, Eaton DH, Vehar GA, Capon DJ, Lawn RM. 1984. Characterisation of the human factor VIII gene. *Nature* 312: 326-30.
- Goodeve AC, Peake IR. 2003. The molecular basis of hemophilia A: genotype-phenotype relationships and inhibitor development. *Semin Thromb Hemost* 29 (1): 23-30.
- Harteveld KL, Losekoot M, Fodde R, Giordano PC, Bernini LF. 1997. The involvement of Alu repeats in recombination events at the alpha-globin gene cluster: characterization of two alpha-zero-thalassemia deletion breakpoints. *Hum Genet* 99: 528-34.
- Kemball-Cook G, Tuddenham EG, Wacey AI. 1998. The factor VIII Structure and Mutation Resource Site: HAMSTeRS version 4. *Nucleic Acids Res* 1: 26 (1): 216-9.
- Krepelova A, Vorlova Z, Zavadil J, Brdika R. 1992. Factor VIII gene deletions in haemophilia A patients in Czechoslovakia. *British J Haematol* 81: 271-276.
- Kricker MC, Drake JW, Radman M. 1992. Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. *Proc Natl Acad Sci USA* 89: 1075-1079.
- Lahiri DK, Nuremberg JI. 1991. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucl Acids Res* 19: 5444.
- Lakich D, Kazazian HH, Antonarakis SE, Gitschier J. 1993. Inversions disrupting the factor VIII gene are common cause of severe Haemophilia A. *Nature Genet* 5: 236-41.
- Lander ES and the International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.
- Lopes J, Tardieu S, Silander K, Blair I, Vanderberghe A, Palau F, Ruberg M, Brice A, LeGern E. 1999. Homologous DNA exchanges in humans can be explained by the yeast double-strand break repair model: a study of 17p11.2 rearrangements associated with CMT1A and HNPP. *Human Molec Genet* 8: 2285-2292.
- Millar DS, Steinbrecher RA, Wieland K, Grundy CB, Martinowitz U, Krawczak M, Zoll B, Whitmore D, Stephenson J, Mibashan RS et al. 1990. The molecular genetic analysis of haemophilia A; characterization of six partial deletions in the factor VIII gene. *Hum Genet* 86: 219-227.
- Moore JK, Haber JE. 1996. Capture of retrotransposon DNA at the sites of chromosomal double strand breaks. *Nature* 383: 644-646.
- Naylor J, Brinke A, Hassock S, Green PM, Giannelli F. 1993. Characteristic mRNA abnormality found in half the patients with severe Haemophilia A is due to large inversions. *Hum Molec Genet* 2: 1773-8.
- Naylor JA, Buck D, Green PM, Williamson H, Bentley D, Giannelli F. 1995. Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions *Human Molec Genet* 4: 1217-24.
- Rüdiger NS, Gregersen N, Kielland-Brandt MC. 1995. One short well conserved region of Alu-sequences is involved in human gene rearrangements and has homology with prokaryotic chi. *Nucleic Acids Res* 23: 256-260.
- Salviato R, Belvini D, Are A, Radossi P, Tagariello G. 2002. Large FVIII gene deletion confers very high risk of inhibitor development in three related severe haemophiliacs. *Haemophilia* 8: 17-21.

## 10 Rossetti et al.

- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning. A laboratory manual. Book 3. Appendix 2.* 2nd ed. Cold Spring Harbor Laboratory Press.
- Shibata M, Shima M, Morichika S, McVey J, Tuddenham EGD, Tanaka I, Suzuki H, Nogami K, Minamoto Y, Hato T, Saenko EL, Scandella D, Yoshioka A. 2000. An alloantibody recognizing the FVIII A1 domain in a patient with CRM reduced haemophilia A due to deletion of a large portion of the A1 domain DNA sequence. *Thromb Haemost* 84: 442-448.
- Sommer SS, Scaringe WA, Hill KA. 2002. Is Alu-mediated recombination an important cause of hemophilia? *Thromb Haemost* 88: 3-4.
- Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. 1983. The double-strand-break repair model for recombination. *Cell* 33: 25-35.
- Tvrđik T, Marcus S, Hou SM, Falt S, Noori P, Podlutska N, Hanefeld F, Stromme P, Lambert B. 1998. Molecular characterization of two deletions events involving Alu-sequences, one novel base substitution and two tentative hotspot mutations in the hypoxanthine phosphoribosyltransferase (HPRT) gene in five patients with Lesch-Nyhan syndrome. *Hum Genet* 103 (3): 311-8
- Van de Water N, Williams R, Ockeford, Browett P. 1998. A 20.7 kb deletion within the factor VIII gene associated with LINE-1 element insertion. *Thromb Haemost* 79: 938-942.
- Vidal F, Farssac E, Tusell J, Puig L, Gallardo D. 2002. First molecular characterization of an unequal homologous Alu-mediated recombination event responsible for haemophilia. *Thromb Haemost* 88: 12-16.
- Viel A, Petronzelli F, Della Puppa L, Lucci-Cordisco E, Fornasarig M, Pucciarelli S, Rovella V, Quaia M, Ponz de Leon M, Boiocchi M, Genuardi M. 2002. Different molecular mechanisms underlie genomic deletions in the MLH1 Gene. *Hum Mutat* 20(5): 368-74.
- Waldman AS, Liskay RM. 1987. Differential effects of base-pair mismatch on intrachromosomal versus extrachromosomal recombination in mouse cells. *Proc Natl Acad Sci USA* 84: 5340-5344.
- Williams IJ, Abuzenadah A, Winship PR, Preston FE, Dolan G, Wright J, Peake IR, Goodeve AC. 1998. Precise carrier diagnosis in families with Haemophilia A: Use of conformation sensitive gel electrophoresis for mutation screening and polymorphism analysis. *Thromb Haemost* 79: 723-6.
- Woods-Samuels P, Kazazian HH Jr, Antonarakis SE. 1991. Nonhomologous recombination in the human genome: deletions in the human factor VIII gene. *Genomics* 10: 94-101.
- Yang D, Waldman AS. 1997. Fine-resolution analysis of products of intrachromosomal homeologous recombination in mammalian cells. *Mol Cell Biol* 17: 3614-3628.
- Youssoufian H, Antonarakis SE, Aronis S, Tsiftis G, Philips DG, Kazazian HH Jr. 1987. Characterization of five partial deletions of the factor VIII gene. *Proc Natl Acad Sci USA* 84: 3772-3776.
- Zucman-Rossi J, Batzer MA, Stoneking M, Delattre O, Thomas G. 1997. Interethnic polymorphism of EWS intron 6: genome plasticity mediated by Alu retroposition and recombination. *Hum Genet* 99: 357-363.