ARTICLE

Next-generation sequencing and recombinant expression characterized aberrant splicing mechanisms and provided correction strategies in factor VII deficiency

Paolo Ferraresi,[1] Dario Balestra,[1] Caroline Guittard,[2] Delphine Buthiau,[2] Brigitte Pan-Petesh,[3] Iva Maestri,[4] Roula Farah,[5] Mirko Pinotti[1] and Muriel Giansily-Blaizot[2]

[1]Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy; [2]Department of Biological Haematology, CHU Montpellier, Université Montpellier, Montpellier, France; [3]Haemophilia Centre, CHU Brest, Brest, France; [4]Department of Morphology, Surgery and Experimental Medicine, University of Ferrara, Ferrara, Italy and [5]Department of Pediatrics, Saint George Hospital University Medical Center, Beirut, Lebanon

Ferrata Storti Foundation

Haematologica 2020
Volume 105(3):829-837

ABSTRACT

Despite the exhaustive screening of *F7* gene exons and exon-intron boundaries and promoter region, a significant proportion of mutated alleles remains unidentified in patients with coagulation factor VII deficiency. Here, we applied next-generation sequencing to 13 FVII-deficient patients displaying genotype-phenotype discrepancies upon conventional sequencing, and identified six rare intronic variants. Computational analysis predicted splicing effects for three of them, which would strengthen (c.571+78G>A; c.806-329G>A) or create (c.572-392C>G) intronic 5' splice sites (5'ss). In *F7* minigene assays, the c.806-329G>A was ineffective while the c.571+78G>A change led to usage of the +79 cryptic 5'ss with only trace levels of correct transcripts (3% of wild-type), in accordance with factor VII activity levels in homozygotes (1-3% of normal). The c.572-392C>G change led to pseudo-exonization and frame-shift, but also substantial levels of correct transcripts (approx. 70%). However, this variant was associated with the common *F7* polymorphic haplotype, predicted to further decrease factor VII levels; this provided some kind of explanation for the 10% factor VII levels in the homozygous patient. Intriguingly, the effect of the c.571+78G>A and c.572-392C>G changes, and particularly of the former (the most severe and well-represented in our cohort), was counteracted by antisense U7snRNA variants targeting the intronic 5'ss, thus demonstrating their pathogenic role. In conclusion, the combination of next-generation sequencing of the entire *F7* gene with the minigene expression studies elucidated the molecular bases of factor VII deficiency in 10 of 13 patients, thus improving diagnosis and genetic counseling. It also provided a potential therapeutic approach based on antisense molecules that has been successfully exploited in other disorders.

Introduction

The inherited deficiency of factor VII (FVII), the crucial enzyme triggering blood coagulation,[1] is the most common of the rare coagulation disorders transmitted in an autosomal recessive manner. The clinical features are highly variable, ranging from severe (i.e. intracranial or gastro-intestinal hemorrhages) to milder (i.e. epistaxis) or asymptomatic forms,[2] and the relationship with plasma FVII activity (FVII:C) levels is often elusive. On the other hand, molecular genetic studies combined with functional assays and recombinant expression investigations, detailing the residual FVII levels associated with *F7* gene mutations, have clearly helped define genotype and coagulation and clinical phenotype relationships,[3-5] with implications for diagnosis, prognosis and counseling.

Over 220 point-mutations,[6] a few large genomic rearrangements, and six com-

Correspondence:

MIRKO PINOTTI
pnm@unife.it

MURIEL GIANSILY-BLAIZOT
m-giansily@chu-montpellier.fr

Received: March 5, 2019.

Accepted: July 2, 2019.

Pre-published: July 4, 2019.

doi:10.3324/haematol.2019.217539

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/105/3/829

mon variants have been identified in the *F7* gene (*http://f7-db.eahad.org/*). Nevertheless, in spite of an exhaustive direct sequencing of *F7* exons and exon-intron junctions and of the proximal promoter region, a significant proportion of defective alleles has still not been identified. The rate of uncharacterized *F7* disease alleles ranges from 2% to 8%[7-10] in Europe, and a similar estimate (7%) was made in India.[11]

In this context, subtle intronic variations outside the routinely sequenced exon-intron boundaries could have a pathological impact by impairing the splicing process. In fact, precise exon definition during RNA processing requires the interplay among several exonic and intronic splicing regulatory elements,[12] which can be altered by nucleotide changes and lead to aberrant splicing.[4,13-17] Various examples of "deep" intronic changes associated with mis-splicing have been reported in human disorders, including those involved in coagulation.[18-20] It is worth noting that RNA splicing can be modulated for different purposes, including the development of new therapies.[14,20-26] In this context, next-generation sequencing (NGS) could represent a powerful tool to characterize gene defects in patients with unknown alleles; however, only a few studies have been conducted in coagulation factor disorders.[27-29]

Here, we investigated 13 patients with FVII deficiency forms that have not been explained by mutations identified by conventional sequencing and used NGS to identify six rare intronic variations that could be causative. Through expression studies, we demonstrated that two of them lead to aberrant splicing, which explained the residual FVII levels in most patients and, intriguingly, that these can be rescued by an antisense-based correction approach.

Methods

Patients

Since 1997, 400 FVII-deficient patients with FVII coagulant activity (FVII:C) levels <30% were referred to our laboratory for genetic analysis through conventional screening. Among them, 13 (Table 1) showed a genetic profile that, considering the identified mutated alleles[10,30-32] and the major *F7* functional polymorphisms (c.-325_-324insCCTATATCCT, A2 allele; p.Arg413Gln change, M2 allele),[33-37] appeared to be incompatible with the reduced FVII levels. The local Institutional Review Board approved the study and patients provided informed consent.

Measurement of FVII levels

FVII:C and FVII antigen levels were determined by the one-stage method[38] and Enzyme-Linked-Immunoabsorbant-Assay (Diagnostica Stago, Asnière sur Seine, France), respectively.

DNA genotyping and next-generation sequencing

Conventional Sanger technology was exploited to sequence the *F7* exons, including the intronic boundaries and the 5' untranslated region. Large rearrangements were ruled out using semi-quantitative multiplex fluorescent-polymerase chain reaction (SQF-PCR) assays.[39] NGS of the *F7* gene was designed to cover the intronic regions with the exception of the highly repetitive GC-rich region in intron 2 (legacy nomenclature, intron 1b). The probe-capture custom design targeted the 14893-pbF7 gene in two parts: chr13:113759000-113761350 and chr13:113764600-113775100 accounting for a total of 12,850 base-pairs with a gap of 3250 bp. DNA library generation was performed using the Custom SureSelectQXT Target Enrichment system (Agilent, Santa Clara, CA, USA) on a MiSeq platform (Illumina, San Diego, CA, USA). The sequencing data were stored in FASTQ format and analyzed using two bio-informatics

Table 1. Features of the investigated FVII deficient patients.

Proband	Origins	FVII:C (%)	FVII:Ag(%)	Conventional sequencing		Polymorphic pattern c.325_324ins/ p.Arg413Gln	NGS sequencing	
				Mutation	Zygosity		Additional nucleotide change	Zygosity
#17	Maghreb countries	3	uk	p.Met358Ile[10]	het	A1A1/M1M1	c.571+78G>A	het
#19	France	2	<5	p.Met1Val[10]	het	A1A2/M1M2	c.571+78G>A	het
#31	France	<1	19	p.Cys162Tyr[10]	het	A1A1/M1M1	c.64+305G>A and c.806-329G>A	het
#113	France	1	66	p.Arg364Trp[30]	het	A1A1/M1M1	c.571+78G>A and c.291+846C>T	het
#262	France	3	uk	p.Gln160Arg[31]	het	A1A1/M1M1	c.806-329G>A	het
#341	France	16	uk	c.430+1G>A[32]	het	A1A1/M1M1	c.64+305G>A and c.681+132G>T	het
#214	France	1	3	p.Arg58ProfsX92	het	A1A1/M1M1	c.571+78G>A	het
#28	France	13	39	p.Arg59Trp	het	A1A1/M1M1	c.571+78G>A	het
#377	Maghreb countries	23	uk			A2A2/M2M2	c.572-392C>G	het
#284 *	Maghreb countries	20	15			A1A2/M1M2	c.571+78G>A	het
#90 *	Lebanon	<1	uk			A1A1/M1M1	c.571+78G>A	hom
#15	Maghreb countries	3	uk			A1A1/M1M1	c.571+78G>A	hom
#330	France	10	uk			A2A2/M2M2	c.572-392C>G	hom

*Consanguinity. NGS: next-generation sequencing; uk: unknown; A1/A2: decanucleotide insertion c.-325_-324insCCTATATCCT promoter polymorphism rs5742910; M1/M2: p.Arg413Gln polymorphism: rs6046; FVII:C: FVII activity; FVII:Ag: FVII antigen; het: heterozygous; hom: homozygous.

pipelines: SEQUENCE Pilot® module SeqNext, version 4.3.1 (JSI medical systems GmbH, Ettenhein, Germany) and SeqOne® (*https://app.seq.one*). The dataset had a mean coverage of over 300x. For each nucleotide of interest, the sequence depth was at least of 30x and the Phred-based quality score above Q30. Regions that did not reach these criteria were sequenced using Sanger technology (primers available upon request); this also applies to each deleterious variation that was independently checked using Sanger sequencing.

Expression studies with *F7* minigenes

The expression vectors for the c.571+78G>A (pIVS6+78A) and c.572-392C>G (pIVS6-392G) variants were created by site-directed mutagenesis of the pIVS6 wild-type (pIVS6-wt), created by cloning the F7 genomic region spanning exon 5 through exon 7 into the pcDNA3 plasmid.[4] Changes were introduced, as previously described,[40] through the overlapping oligonucleotides $^{5'}$GAAGCAGATCAAAAGTAAGCATGGGATC$^{3'}$ and $^{5'}$GATCCCATGCTTACTTTTGATCTGCTTC$^{3'}$ for the c.572-392C>G mutation. For the c.571+78G>A mutation, we exploited the non-overlapping oligonucleotides $^{5'}$CTGGACAAAAGACAGGTGGG AGTGGC$^{3'}$ and $^{5'}$TAAGATAATCCGTAGTGGGACAGGG ACT$^{3'}$ in a slightly modified protocols that implies, after PCR cycles, the addition of T4 polynucleotide kinase and T4 DNA ligase to ensure circulation of the products.

The U7smOPT expression vectors (pU7smOPT) were created as previously described[41] using a standard SP6 reverse oligonucleotide $^{5'}$ATTTAGGTGACACTATAG$^{3'}$ and the forward mutagenic oligonucleotides $^{5'}$ACAGAGGCCTTTCCGCAcccacctgtcttttgtccaAATTTTTTGGAG$^{3'}$ (U7+78A), $^{5'}$ACAGAGGCCTTTCCGCAtgaagccactcccacctgAATTTTTGGAG$^{3'}$ (U7+78Ash) and $^{5'}$ACAGAGGCCTTTCCGCAcatgcttacttttgatctAATTTTTGGAG$^{3'}$(U7-392G).

One microgram of the pIVS6 variant, alone or with equimolar amounts of the pU7smOPT variant, was transiently transfected in human embryonic kidney cells (HEK293T) by lipofection in 12-well plates.[13] RNA isolation and reverse transcription[42] were followed by PCR using the forward oligonucleotide *T7bisF* ($^{5'}$CACTGCTTACTGGCTTATCGAAAT$^{3'}$, in the pcDNA3 T7 region), either unmodified or $^{5'}$fluorescently labeled (*T7bisF*^{FAM}), and the reverse oligonucleotide *F7ex7R* ($^{5'}$CACAACTGAGCTCCATTCACCAACA$^{3'}$, in exon 7) or *F7PsExR* ($^{5'}$TTCAATCAAGGTCTTGGGCC$^{3'}$, in the pseudo-exon 5b).

Results

Genotyping of FVII deficient patients

Among the 13 selected FVII-deficient patients shown in Table 1, ten had FVII:C levels below 15% and were expected to have two *F7* pathogenic alleles. However, the conventional sequencing did not reveal any *F7* pathogenic allele for patients #15, #90 and #330, and only one for the remaining seven patients (#17, #19, #28, #31, #113, #214, #262). On the other hand, three patients (#284, #341, #377) presented with FVII:C levels between 15% and 30% but displayed only the c.430+1G>A mutation (#341) or the A2M2 polymorphic haplotype (#284, #377), which points towards the presence of an additional *F7* pathogenic allele for each patient. In this scenario, we had a total of 16 *F7* uncharacterized alleles to be explored using NGS.

As far as the clinical phenotype is concerned, 3 of 13 patients were symptomatic. Patient #19 presented with bruises and frequent epistaxis, patient #90 had post-traumatic oral bleeding, spontaneous hematuria and rectal bleeding, and patient #214 suffered from provoked hematoma and severe menorrhagia resolved by replacement therapy.

Next-generation sequencing, besides confirming the presence of the causative variants identified by the Sanger approach, also revealed several deep intronic substitutions. Among them, only those with a coverage of 30x and observed in databases with a minor allele frequency (MAF) <0.05 were analyzed further. Six deep intronic substitutions matched these criteria (Table 2). The c.571+78G>A change, whose pathogenic effect is supported by its co-segregation with the disease phenotype in the family pedigree of patient #28 (Figure 1), was the most frequent in our series, being present in ten alleles from unrelated patients living in various areas, including France, North Africa and Lebanon (Table 1). NGS data prompted us to analyze an enlarged panel of polymorphic deep intronic variants in both c.571+78G>A homozygotes (#15, #90), who were homozygous for the major A1 and M1 polymorphic alleles. However, they differed on other intronic variants. Patient #15, of Tunisian origin, was homozygous for two variants, c.-402A (rs510317) and c.292-672G (rs12431329), that are quite rare, with a minor allele frequency (MAF) of 0.233 and 0.213, respectively. By contrast, patient #90, living in Lebanon, showed the c.-402G and the c.292-672A variants, and displayed two additional deep intronic polymorphic variants in the homozygous state, the c.64+196G>A (rs2774030) and the c.131-394T>C (rs1745939), with a global frequency of the c.64+196A and c.131-394C alleles of 0.551 and 0.739, respectively. Thus, two different haplotypes associated with the c.571+78G>A mutation could be defined: c.-402A, A1 c.64+196G, c.131-394T, c.292-672G, c.571+78A, M1 (Haplotype 1) and c.-402G, A1,

Table 2. Deep intronic mutations found by next-generation sequencing screening.

Change	Intron	Prediction	Score (wt/mutated)	Position	rs	MAF
c.64+305G>A	1a	Weakening cryptic 3'ss	0.31/0.13	-8	36208414	0.005
c.291+846C>T	2	Strengthening cryptic 3'ss	0.51/0.71	-7	565185989	0.004
c.571+78G>A	5	Strengthening cryptic 5'ss	0.11/0.79	-2	764741909	none
c.572-392C>G	5	Creation of new 5'ss	nd/0.98	+1	none	none
c.681+132G>T	6	Strengthening cryptic 3'ss	0.02/0.12	-8	752129277	none
c.806-329G>A	7	Strengthening cryptic 5'ss	0.79/0.99	+3	none	none

Position of the point mutation is referred to the 5' splice site (5'ss) or 3'ss. RS: reference SNP ID number; MAF: minor allele frequency, based on 1000Genome project (*http://www.internationalgenome.org*). *F7* gene reference sequence is NG_009262.1. NNSPLICE 0.9 software (*www.fruitfly.org/seq_tools/splice.html*) was used to predict and calculate the 5'ss or 3'ss of the score. Introns are indicated by legacy nomenclature. wt: wild-type.

c.64+196A, c.131-394C, c.292-672A, c.571+78A, M1 (Haplotype 2). The analysis was also extended to the c.571+78G>A heterozygotes, which revealed that haplotype 1 is compatible with patients of Maghreb origin whereas haplotype 2 is compatible with patients of European origin.

The c.572-392C>G, c.64+305G>A and c.806-329A variants were also relatively frequent as they were found in three, two and two alleles, respectively, whereas the remaining were identified only once.

Computational analysis of splicing regulatory elements

We performed an *in silico* analysis of the six deep-intronic mutations to infer a pathogenic effect on splicing. In principle the nucleotide changes could affect Intronic Splicing regulatory elements such as enhancers (ISE) or Silencers (ISS)[43] or create/strengthen 5' or 3' splice sites (ss). However, regulatory elements generally reside within the first 200 bp of the intron[44] and the main bioinformatics tools (i.e. Human Splicing Finder, *www.umd.be/HSF/*) have been developed to predict exonic elements, which confers prediction of the impact of the investigated changes with an unacceptable degree of speculation. Therefore, we focused the analysis on 5'ss and 3'ss (*www.fruitfly.org/seq_tools/splice.html*), which predicted that the c.64+305G>A, c.291+846C>T, c.681+132G>T nucleotide changes do not appreciably strengthen cryptic splice sites. Concerning the c.571+78G>A, c.572-392C>G and c.806-329G>A variants, the introduction of the nucleotide changes would result in the creation (c.572-392C>G) or remarkable strengthening (c.571+78G>A and c.806-329G>A) of a cryptic 5'ss (Table 2).

In vitro characterization of the splicing variants

Based on the bioinformatics prediction of the impact of these variants on splicing, on the number of affected alleles, on the MAF, and on identification in homozygous conditions, we selected the c.571+78G>A, c.572-392C>G and c.806-329G>A changes for further characterization. Due to the impossibility of investigating *F7* mRNA processing in patients' hepatocytes, the physiological site of *F7* synthesis, or to the unavailability of fresh leukocytes as ectopic source of *F7* mRNA, we exploited the expression of *F7* minigenes (Figure 2A). The transfection of the pIVS6-wt minigene and splicing pattern analysis revealed correct splicing (Figure 2C, transcript 2) but also, albeit to a lesser extent, exon 6 skipping (transcript 1) and usage of the weak cryptic intronic 5'ss at position +79 that leads to partial intron retention (transcript 3G).

The splicing pattern analysis of cells transfected with the pIVS6+78A minigene showed an aberrant transcript (Figure 2C, transcript 3A) that, upon sequencing, indicated the usage of the strengthened intronic 5'ss at position +79 (Figure 2B, transcript 3A). This leads to partial intron retention resulting in a deleted and frame-shifted mRNA harboring a premature nonsense triplet at position p.201, not expected to produce a functional FVII protein. To evaluate the presence of residual FVII levels, the RT-PCR was fluorescently labeled and the amplicons evaluated by denaturing capillary electrophoresis, which ensures high sensitivity. This approach led us to identify very low levels of correct transcripts, which roughly accounted for approximately 3% of the overall transcripts (Figure 3B).

While the splicing analysis of c.806-329G>A construct in pCDNA3 did not reveal any alteration by transient transfection in HEK293T (*Online Supplementary Appendix* and *Online Supplementary Figure S1*) or in Baby Hamster Kidney cells (*data not shown*), the assessment of splicing pattern of cells transfected with the FVII minigene harboring the c.572-392C>G change revealed splicing abnormalities (Figure 2B). In addition to the correctly spliced mRNA, we identified transcripts arising from skipping of

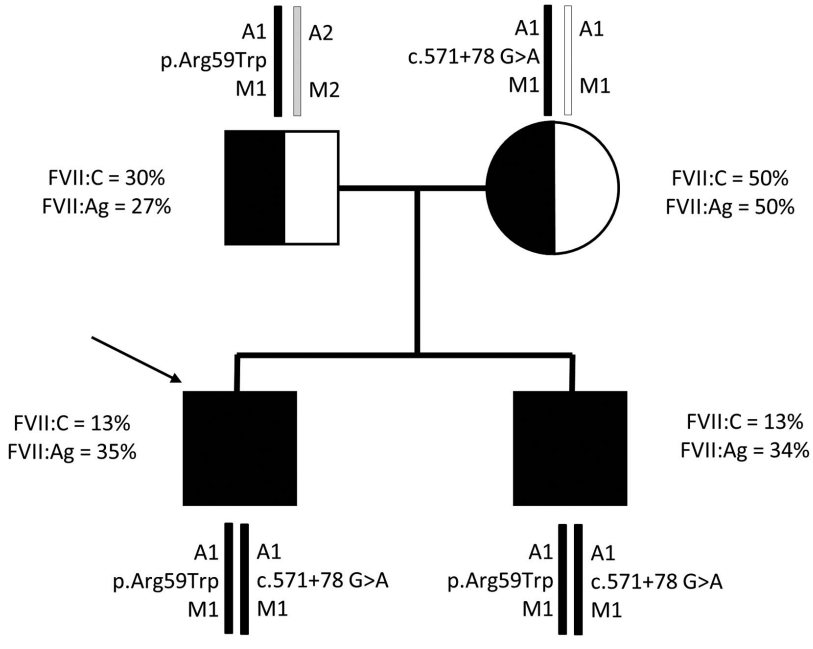

Figure 1. Family pedigree of patient #28. The propositus is indicated by an arrow. Mutation, nucleotide changes, A1A2 (rs5742910), M1M2 (p.Arg413Gln, rs6046) haplotypes, coagulant factor VII (FVII:C) and factor VII antigen (FVII:Ag) are indicated.

exon 6, from usage of the weak cryptic 5'ss at position +79 and, most importantly, from a pseudo-exonization event (transcripts 1, 3G, and 4, respectively). More precisely, this new pseudo-exon 5b originates from the usage of a cryptic 5'ss and of a 282 bp upstream cryptic 3'ss. This finding was further strengthened by a PCR using a primer in the pseudo-exon 5b (F7PsExR), which gave rise to amplified fragments only in cells expressing the pIVS6-392G construct (Figure 2B, lower panel). The inclusion of the pseudo-exon leads to a frame-shifted mRNA with a premature nonsense codon, predicted to encode a dysfunctional FVII protein. Semi-quantitative evaluation of transcripts by fluorescent labeling of amplicons and denaturing capillary electrophoresis revealed that the correct and the aberrant forms are present in the relative proportion of 74% and 26%, respectively (Figure 3C), compared to 90% and 10% in the pIVS6-wt context (Figure 3A).

Overall, these data demonstrate that both mutations exert their detrimental effect by impairing FVII splicing, strengthening the usage of cryptic 5'ss.

Investigation by antisense U7snRNA

Since the observed aberrant splicing is caused by the usage of new 5'ss, we hypothesized that masking them would weaken or abolish their detrimental role. To this purpose, we exploited variants of the U7 small nuclear RNA (U7smOPT)[41] as potent antisense molecules to target the alternative splice sites (Figure 4A).

Co-expression of the c.571+78G>A change with antisense U7smOPT variants resulted in an appreciable rescue of splicing, as evaluated by densitometric analysis of bands upon semi-quantitative PCR. In particular, the proportion of correct transcripts, barely appreciable in untreated conditions, remarkably increased to approximately 9% or to approximately 20% of total transcripts upon co-expression of the pU7+78Ash or pU7+78A, respectively (Figure 4B, bottom).

Concerning the c.572-392G variant, co-expression of pU7-392G, designed on the cryptic 5'ss, resulted in a 3-fold reduction in aberrantly spliced mRNA containing the pseudo-exon 5b and conversely favored (1.3-fold increased) the synthesis of correctly spliced transcripts that rose from approximately 70% to approximately 90% of all forms (Figure 4B, bottom), resembling the proportion observed in the wild-type context.

Overall, these data further demonstrate the causative role of the mutations that create/strengthen cryptic intronic 5'ss.

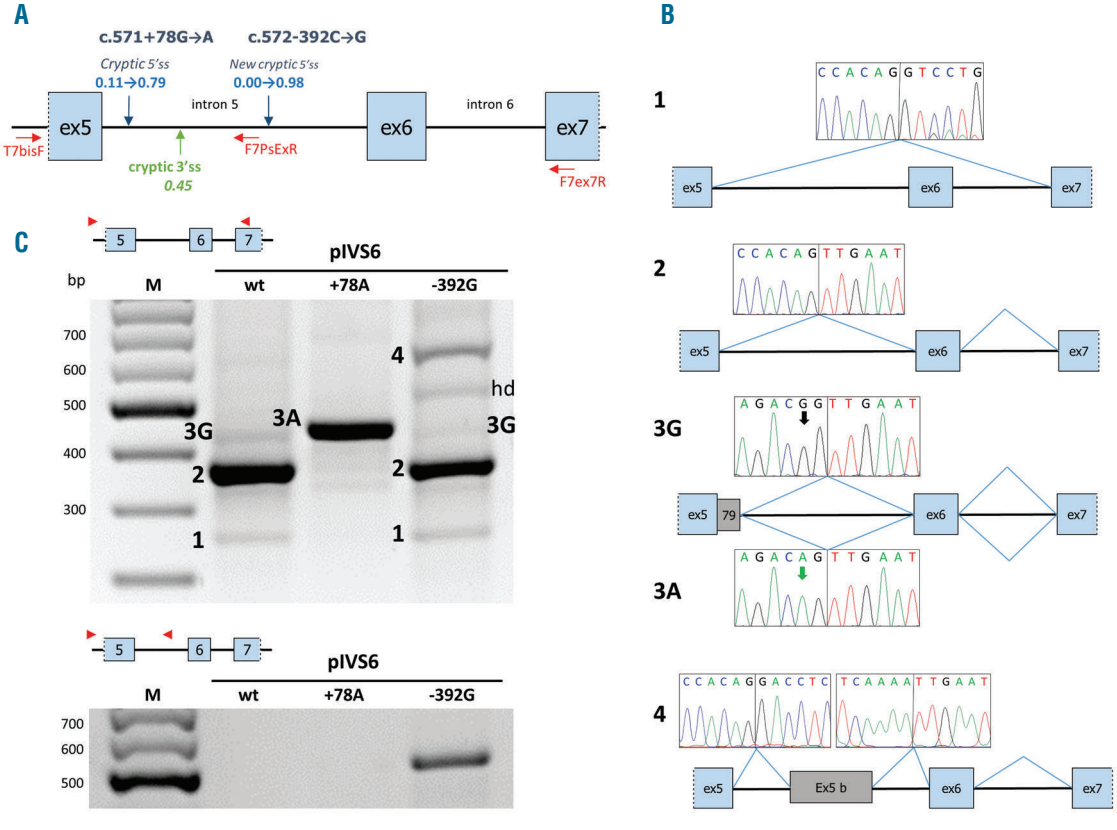

Figure 2. Alternative splicing patterns associated with the c.571+78G>A and c.572-392C>G mutations. (A) Schematic representation of the pIVS6 minigene. Mutations (blue) are reported on top. The presence of cryptic splice sites (5'ss in light blue, 3'ss in green), with related scores are indicated by arrows. Polymerase chain reaction (PCR) oligonucleotides are shown in red. (B) The schematic representation of splicing patterns is reported together with relative sequencing chromatograms of the amplicons, obtained by T7bisF-F7ex7R PCR (see panel A). PCR fragments were cloned before sequencing. (C) Splicing pattern analysis in HEK293T cells transiently transfected with pIVS6 wild type (wt) or with pIVS6 variants c.571+78A (+78A) and c.572-392G (-392G); the PCR with T7bisF-F7ex7R oligonucleotides is reported in the upper panel; PCR with T7bisF-F7psExR oligonucleotides specifically designed to amplify transcripts with the pseudo-exon5b, is reported in the lower panel. M:100 bp ladder; hd: heteroduplex.

Discussion

Uncharacterized *F7* pathogenic alleles are mentioned in all patient databases, and NGS would represent a powerful tool to tackle this; however, so far, this has not been fully explored. Here, we applied this approach in 13 FVII deficient patients who were only partially characterized through conventional sequencing, and identified a panel of deep intronic substitutions as candidates to explain the reduced FVII:C levels in patients. However, as for the numerous deep intronic nucleotide changes identified by NGS and associated with inherited diseases, their pathogenic role requires experimental support.

The evidence for the pathogenicity of deep intronic variations relies on several clinical and molecular observations. The first aspect to consider is their virtual absence in databases. This led us to exclude from our selection the c.64+305G>A and c.291+846C>T variants, with a minor allelic frequency of 0.005 and 0.004, respectively. The same applies to the c.681+132G>T change, reported in dbSnp databases as rs752129277 but with an estimated frequency of <0.0004. Another aspect that could suggest the pathogenicity of the nucleotide changes is their distribution among affected relatives or non-related individuals. This was the case of the c.571+78G>A, c.572-392C>G and c.806-329G>A changes that occurred in eight, two and two unrelated FVII-deficient patients, respectively. Altogether these elements prompted us to explore the impact of these three variants on the splicing process through the expression of minigenes in eukaryotic cells, a well-proven approach used to dissect splicing abnormalities.[4,13,14,42,45]

Regarding the c.571+78G>A variant, the *in vitro* characterization demonstrated an aberrant splicing profile that was consistent with the FVII:C levels reported in patients. Interestingly, the amount of correctly spliced transcripts (approx. 3%) reflects the FVII:C levels (3% and <1%) in the two c.571+78A homozygotes (patients #15 and #90) (Table 1), which explains the asymptomatic or moderate clinical phenotypes. This finding is also consistent with the observation that the mutation co-segregated with the disease phenotype through the pedigree of patient #28, and the 50% FVII:C levels detected in the heterozygous mother.

It is interesting to note that the c.571+78G>A mutation, albeit absent from databases, occurred in eight apparently unrelated patients in our cohort. The polymorphic analysis in the homozygous patients led us to identify two different haplotypes that suggested two distinct mutational

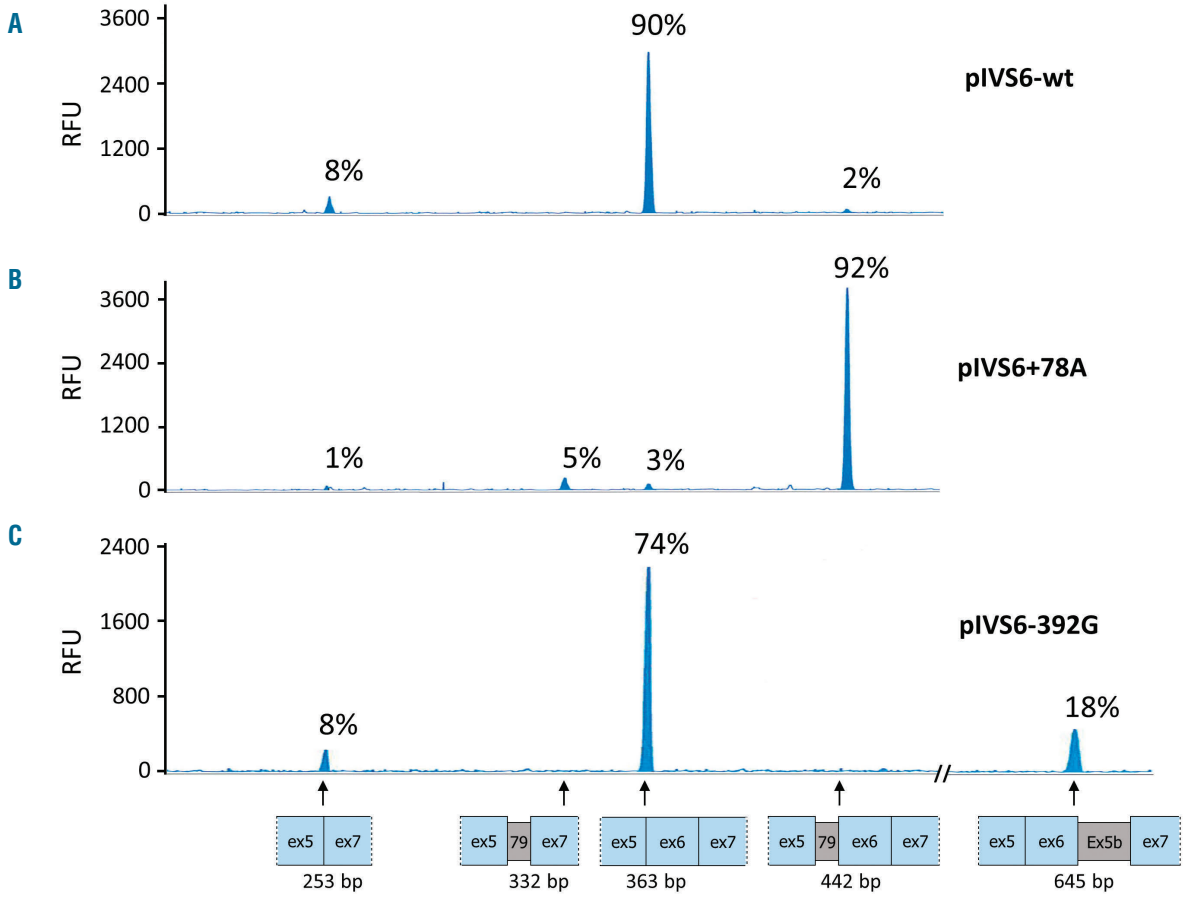

Figure 3. Alternative splicing patterns evaluated by denaturing capillary electrophoresis. Splicing patterns of the pIVS6-wt (A), pIVS6+78A (B) and pIVS6-392G (C) minigenes upon transient transfection in HEK293T cells and evaluated by polymerase chain reaction with T7bisF^{Fam}-F7ex7R oligonucleotides performed at 28 cycles followed by denaturing capillary electrophoresis. The schematic representation of transcripts (not in scale) is reported below and the amplicons base pairs (bp) are indicated at the bottom. The relative amount of transcripts is indicated by percentages. RFU: relative fluorescence units.

events leading to two subsequent founder effects. In particular, one haplotype was found in patients of Maghreb origin whereas the second one was compatible with genotypes of patients of European and Lebanese origins.

Concerning the c.572-392C>G change, the aberrant splicing profile displayed approximately 70% of correctly spliced transcripts, which is not apparently consistent with the FVII:C levels (10%) observed in homozygous patient #330. However, the mutation was associated with the less frequent A2 and M2 polymorphic alleles, which have been demonstrated to halve the FVII expression when present in the homozygous state.[33-37]

The c.571+78G>A and c.572-392C>G mutations were also found in patients with moderately reduced FVII levels (#28, #377, #284). In patients #377 and #284, this can be explained by the fact that: i) the mutations are present in heterozygous condition; and ii) by the additional contribution of the functional polymorphisms A2 and M2. For patient #28, heterozygous for the c.571+78G>A mutation, the residual expression could arise from the allele bearing the p.Arg59Trp change, which, however, has never been characterized.

In contrast to the c.571+78G>A and c.572-392C>G variants, the *in vitro* splicing analysis did not reveal detrimental effects on splicing for the c.806-329G>A change. Therefore, we were unable to explain the genotype-phenotype relationship for three patients (#31, #262 and #341) who were not carriers of the c.571+78G>A or c.572-392C>G mutations. Since we have not identified other candidate pathogenic variants besides the mutations previously identified by conventional sequencing, it is tempting to speculate that the genetic defect could be in the unexplored highly repetitive rich GC region of intron 2 or in the 5' or 3' regulatory regions of the *F7* gene.

Knowledge of the alternative splicing patterns and of the mechanisms involved offers the opportunity to design correction strategies that could have therapeutic implications.[20-22] Here, we exploited variants of the U7 small nuclear RNA, the RNA component of the U7 small nuclear ribonucleoprotein that is biologically involved in histone RNA 3' end processing.[46] By changing the Sm consensus sequence of the endogenous U7snRNA, it is possible to express U7snRNA variants (U7smOPT) that no longer modulate histone processing, but can bind to RNA targets though base-pair interaction and efficiently accumulate into nucleus as snRNP.[41] Therefore, by changing the 5' tail of the U7smOPT, it is possible to target a desired RNA sequence and avoid its recognition by splic-

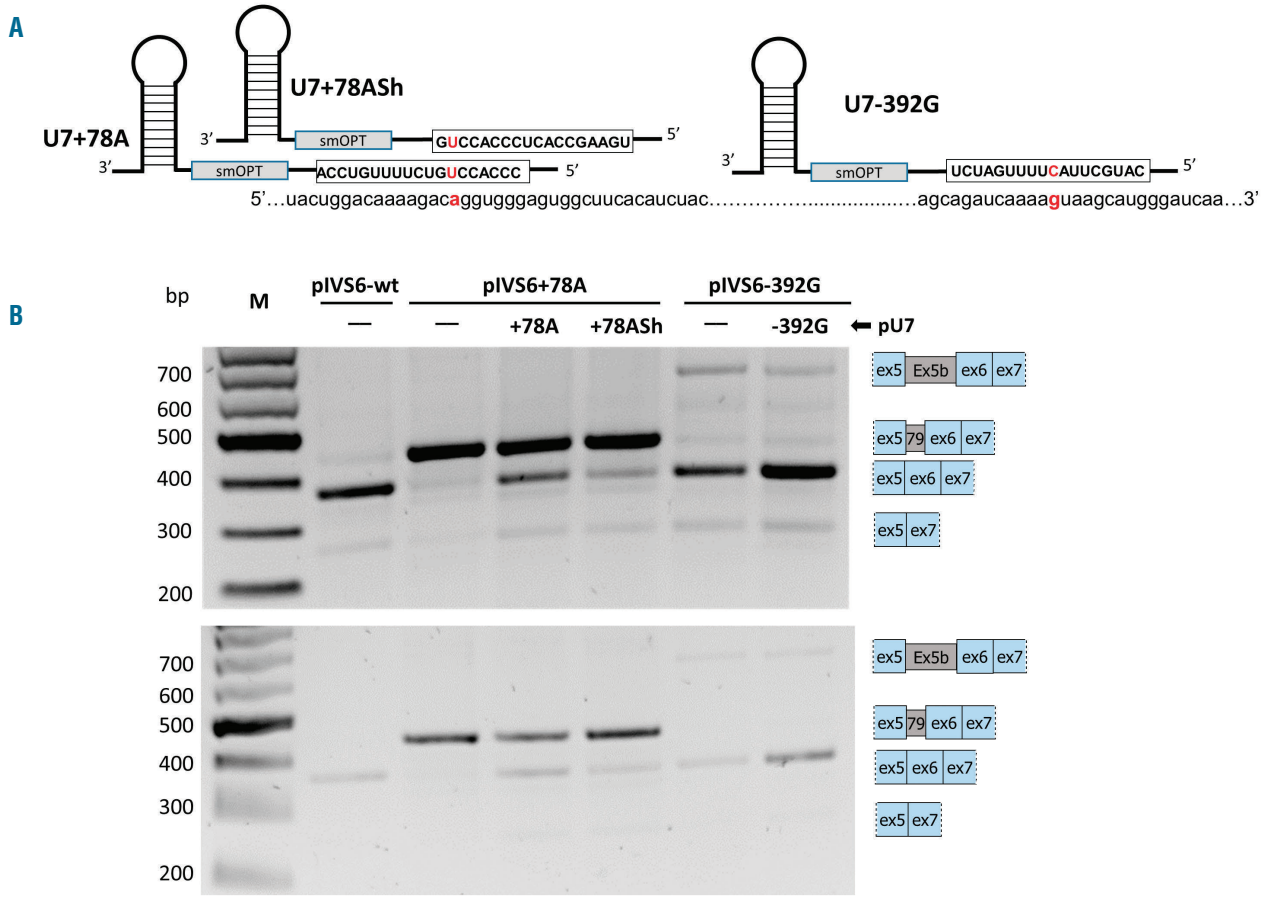

Figure 4. Investigation of aberrant splicing mechanisms by using antisense U7smOPT variants. (A) Schematic representation of engineered U7smOPT exploited in this study. The sequence of intronic mRNA and of the engineered 5'tail of U7smOPT with relative base-pairing is reported. The nucleotide changes identified by next-generation sequencing are indicated in bold and red, as well as the corresponding base in U7smOPT antisense sequence. (B) Splicing pattern analysis in HEK293T cells transiently transfected with pIVS6 wild type (wt) and variants (c.571+78A and c.572-392G) alone or in combination with engineered U7smOPT (pU7). PCR with T7bisF-F7ex7R oligonucleotides performed at 32 cycles (top) or, for semi-quantitative evaluation, at 25 cycles (bottom). M: 100 bp ladder.

ing factors. This opportunity has been exploited to induce both exon skipping for therapeutic purposes[47,48] and to dissect splicing regulatory elements by masking them.[20,22,49] In our study, the observation that antisense U7smOPT variants masking the cryptic 5'ss were able to rescue the splicing pattern further demonstrated the pathogenic role of the nucleotide changes and provided a correction approach that has potential therapeutic implications for the c.571+78A mutation. In fact, if translated into patients, the correct transcripts rescued for this mutation (a well-represented change in our patient cohort and associated with severe forms) would account for FVII levels well beyond the therapeutic threshold.

In conclusion, the combination of NGS of the entire F7 gene with the expression of minigenes elucidated the molecular bases of FVII deficiency in ten out of thirteen FVII deficient patients, thus improving diagnosis and genetic counselling, and provided insight into a potential therapeutic approach based on antisense technology, successfully exploited in other disorders.

Acknowledgments

The authors would like to thank Dr D. Schümperli for sharing the U7-SmOPT construct. This research was supported by grants from Bayer (2017-INT.A-BD_001 to D.Balestra), the University of Ferrara and the University of Montpellier. The authors gratefully acknowledge H. Chambost, K. Pouymayou (Marseille), S. Clayssens (Toulouse), R. D'Oiron, I. Martin-Toutain (Paris), V. Gay (Chambery), Y. Chevalier, S Le-Quellec, S. Meunier, L. Rugeri (Lyon), C. Montaud (Pau) and R. Navarro (Montpellier, France) for sending samples and L.Bison for technical assistance.

References

1. Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation, maintenance, and regulation. Biochemistry. 1991; 30(43):10363-10370.
2. Bernardi F, Dolce A, Pinotti M, et al. Major differences in bleeding symptoms between factor VII deficiency and hemophilia B. J Thromb Haemost. 2009;7(5):774-779.
3. Pinotti M, Etro D, Bindini D, et al. Residual factor VII activity and different hemorrhagic phenotypes in CRM(+) factor VII deficiencies (Gly331Ser and Gly283Ser). Blood. 2002;99(4):1495-1497.
4. Cavallari N, Balestra D, Branchini A, et al. Activation of a cryptic splice site in a potentially lethal coagulation defect accounts for a functional protein variant. Biochim Biophys Acta - Mol Basis Dis. 2012;1822(7):1109-1113.
5. Branchini A, Ferrarese M, Lombardi S, Mari R, Bernardi F, Pinotti M. Differential functional readthrough over homozygous nonsense mutations contributes to the bleeding phenotype in coagulation factor VII deficiency. J Thromb Haemost. 2016;14(10):1994-2000.
6. McVey JH, Boswell E, Mumford AD, Kemball-Cook G, Tuddenham EG. Factor VII deficiency and the FVII mutation database. Hum Mutat. 2001;17(1):3-17.
7. Herrmann FH, Wulff K, Auerswald G, et al. Factor VII deficiency: clinical manifestation of 717 subjects from Europe and Latin America with mutations in the factor 7 gene. Haemophilia. 2009;15(1):267-280.
8. Millar DS, Kemball-Cook G, McVey JH, et al. Molecular analysis of the genotype-phenotype relationship in factor VII deficiency. Hum Genet. 2000;107(4):327-342.
9. Peyvandi F, Jenkins P V, Mannucci PM, et al. Molecular characterisation and three-dimensional structural analysis of mutations in 21 unrelated families with inherited factor VII deficiency. Thromb Haemost. 2000; 84(2):250-257.
10. Giansily-Blaizot M, Aguilar-Martinez P, Biron-Andreani C, Jeanjean P, Igual H, Schved JF. Analysis of the genotypes and phenotypes of 37 unrelated patients with inherited factor VII deficiency. Eur J Hum Genet. 2001;9(2):105-112.
11. Mota L, Shetty S, Idicula-Thomas S, Ghosh K. Phenotypic and genotypic characterization of Factor VII deficiency patients from Western India. Clin Chim Acta. 2009;409(1-2):106-111.
12. De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. Wiley Interdiscip Rev RNA. 2013;4(1):49-60.
13. Pinotti M, Toso R, Redaelli R, Berrettini M, Marchetti G, Bernardi F. Molecular mechanisms of FVII deficiency: expression of mutations clustered in the IVS7 donor splice site of factor VII gene. Blood. 1998; 92(5):1646-1651.
14. Balestra D, Barbon E, Scalet D, et al. Regulation of a strong F9 cryptic 5'ss by intrinsic elements and by combination of tailored U1snRNAs with antisense oligonucleotides. Hum Mol Genet. 2015; 24(17):4809-4816.
15. Nuzzo F, Bulato C, Nielsen BI, et al. Characterization of an apparently synonymous F5 mutation causing aberrant splicing and factor V deficiency. Haemophilia. 2015; 21(2):241-248.
16. Donadon I, McVey JH, Garagiola I, et al. Clustered F8 missense mutations cause hemophilia A by combined alteration of splicing and protein biosynthesis and activity. Haematologica. 2018;103(2):344-350.
17. Scalet D, Sacchetto C, Bernardi F, Pinotti M, Van De Graaf SFJ, Balestra D. The somatic FAH C.1061C>A change counteracts the frequent FAH c.1062+5G>A mutation and permits U1snRNA-based splicing correction. J Hum Genet. 2018;63(5):683-686.
18. Castoldi E, Duckers C, Radu C, et al. Homozygous F5 deep-intronic splicing mutation resulting in severe factor V deficiency and undetectable thrombin generation in platelet-rich plasma. J Thromb Haemost. 2011;9(5):959-968.
19. Castaman G, Giacomelli SH, Mancuso ME, et al. Deep intronic variations may cause mild hemophilia A. J Thromb Haemost. 2011;9(8):1541-1548.
20. Nuzzo F, Radu C, Baralle M, et al. Antisense-based RNA therapy of factor V deficiency: In vitro and ex vivo rescue of a F5 deep-intronic splicing mutation. Blood. 2013; 122(23):3825-3831.
21. Goyenvalle A, Vulin A, Fougerousse F, et al. Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. Science. 2004;306(5702):1796-1799.
22. Meyer K, Marquis J, Trüb J, et al. Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation. Hum Mol Genet. 2009; 18(3):546-555.
23. Alanis EF, Pinotti M, Mas AD, et al. An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. Hum Mol Genet. 2012;21(11):2389-2398.
24. Balestra D, Scalet D, Pagani F, et al. An exon-specific U1snRNA induces a robust factor IX activity in mice expressing multiple human FIX splicing mutants. Mol Ther - Nucleic Acids. 2016;5(10):e370.
25. Scalet D, Balestra D, Rohban S, et al. Exploring splicing-switching molecules for Seckel syndrome therapy. Biochim Biophys Acta - Mol Basis Dis. 2017;1863(1):15-20.
26. Scalet D, Maestri I, Branchini A, Bernardi F, Pinotti M, Balestra D. Disease-causing variants of the conserved +2T of 5′ splice sites can be rescued by engineered U1snRNAs. Hum Mutat. 2018;(October 2018):48-52.
27. Pezeshkpoor B, Zimmer N, Marquardt N, et al. Deep intronic "mutations" cause hemophilia A: Application of next generation sequencing in patients without detectable mutation in F8 cDNA. J Thromb Haemost. 2013;11(9):1679-1687.
28. Bastida JM, del Rey M, Lozano ML, et al. Design and application of a 23-gene panel by next-generation sequencing for inherited coagulation bleeding disorders. Haemophilia. 2016;22(4):590-597.
29. Simeoni I, Stephens JC, Hu F, et al. A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders. Blood. 2016;127(23):2791-2803.
30. Matsushita T, Kojima T, Emi N, Takahashi I, Saito H. Impaired human tissue factor-mediated activity in blood clotting factor VIINagoya (Arg304-->Trp). Evidence that a region in the catalytic domain of factor VII is important for the association with tissue factor. J Biol Chem. 1994;269(10):7355-7363.
31. Kavlie A, Orning L, Grindflek A, Stormorken H, Prydz H. Characterization of a factor VII molecule carrying a mutation in the second epidermal growth factor-like domain. Thromb Haemost. 1998;79(6):1136-1143.
32. Arbini AA, Mannucci M, Bauer KA. A Thr359Met mutation in factor VII of a patient with a hereditary deficiency causes defective secretion of the molecule. Blood. 1996;87(12):5085-5094.
33. Bernardi F, Marchetti G, Pinotti M, et al.

Factor VII gene polymorphisms contribute about one third of the factor VII level variation in plasma. Arterioscler Thromb Vasc Biol. 1996;16(1):72-76.
34. Girelli D, Russo C, Ferraresi P, et al. Polymorphisms in the factor VII gene and the risk of myocardial infarction in patients with coronary artery disease. N Engl J Med. 2000;343(11):774-780.
35. Kudaravalli R, Tidd T, Pinotti M, et al. Polymorphic changes in the 5ʹʹʹ flanking region of Factor VII have a combined effect on promoter strength. Thromb Haemost. 2002;88(5):763-767.
36. Bozzini C, Girelli D, Bernardi F, et al. Influence of polymorphisms in the factor VII gene promoter on activated factor VII levels and on the risk of myocardial infarction in advanced coronary atherosclerosis. Thromb Haemost. 2004;92(3):541-549.
37. Sabater-Lleal M, Chillón M, Howard TE, et al. Functional analysis of the genetic variability in the F7 gene promoter. Atherosclerosis. 2007;195(2):262-268.
38. Giansily-Blaizot M, Verdier R, Biron-Adréani C, et al. Analysis of biological phenotypes from 42 patients with inherited factor VII deficiency: Can biological tests predict the bleeding riskʔ Haematologica. 2004;89(6):704-709.
39. Giansily-Blaizot M, Thorel D, Khau Van Kien P, et al. Characterisation of a large complex intragenic re-arrangement in the FVII gene (F7) avoiding misdiagnosis in inherited factor VII deficiency. Br J Haematol. 2007; 138(3):359-365.
40. Ferrarese M, Testa MF, Balestra D, Bernardi F, Pinotti M, Branchini A. Secretion of wild-type factor IX upon readthrough over F9 pre-peptide nonsense mutations causing hemophilia B. Hum Mutat. 2018;39(5):702–708.
41. Meyer K, Schümperli D. Antisense Derivatives of U7 Small Nuclear RNA as Modulators of Pre-mRNA Splicing. In: Alternative Pre–mRNA Splicing. Wiley-Blackwell; 2012:481-494.
42. Pinotti M, Rizzotto L, Balestra D, et al. U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency. Blood. 2008;111(5):2681-2684.
43. Yeo GW, Van Nostrand EL, Liang TY. Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. PLoS Genet. 2007;3(5):e85.
44. Majewski J, Ott J. Distribution and characterization of regulatory elements in the human genome. Genome Res. 2002; 12(12):1827-1836.
45. Pinotti M, Toso R, Girelli D, et al. Modulation of factor VII levels by intron 7 polymorphisms: population and in vitro studies. Blood. 2000;95(11):3423-3428.
46. Dominski Z, Marzluff WF. Formation of the 3' end of histone mRNA. Gene. 1999; 239(1):1-14.
47. Brun C, Suter D, Pauli C, et al. U7 snRNAs induce correction of mutated dystrophin pre-mRNA by exon skipping. Cell Mol Life Sci. 2003;60(3):557-566.
48. Suter D, Tomasini R, Reber U, Gorman L, Kole R, Schümperli D. Double-target antisense U7 snRNAs promote efficient skipping of an aberrant exon in three human β-thalassemic mutations. Hum Mol Genet. 1999;8(13):2415-2423.
49. Gorman L, Suter D, Emerick V, Schümperli D, Kole R. Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. Proc Natl Acad Sci U S A. 1998;95(9):4929-4934.