Journal of Thrombosis and Haemostasis, **11**: 1679–1687 DOI: 10.1111/jth.12339

ORIGINAL ARTICLE

Deep intronic 'mutations' cause hemophilia A: application of next generation sequencing in patients without detectable mutation in *F8* cDNA

B. PEZESHKPOOR,* N. ZIMMER,* N. MARQUARDT,* I. NANDA,† T. HAAF,† U. BUDDE,‡ J. OLDENBURG* and O. EL-MAARRI*

**Institute of Experimental Hematology and Transfusion Medicine, University of Bonn, Bonn; †Institute of Human Genetics, University of Würzburg; and ‡Medilys Laborgesellschaft, Hamburg, Germany*

To cite this article: Pezeshkpoor B, Zimmer N, Marquardt N, Nanda I, Haaf T, Budde U, Oldenburg J, El-Maarri O. Deep intronic 'mutations' cause hemophilia A: application of next generation sequencing in patients without detectable mutation in *F8* cDNA. *J Thromb Haemost* 2013; **11**: 1679–87.

Summary. *Background:* In a small group of typical hemophilia A (HA) patients no mutations in the *F8* coding sequence (cDNA) could be found. In the current study, we performed a systematic screening of genetic and non-genetic parameters associated with reduced FVIII:C levels in a group of mostly mild HA (only one moderate) patients with no detectable mutations in *F8* cDNA. *Methods:* We determined FVIII and VWF activity and antigen levels and performed VWF-FVIII binding (VWF:FVIIIB) and VWF-collagen binding assays (VWF:CB) as well as VWF multimer analysis. *VWF* was completely sequenced to exclude mutations. The *F8* locus, including the introns, was sequenced using overlapping long-range PCRs (LR-PCRs) combined with a next generation sequencing (NGS) approach. Moreover, the *F8* mRNA was analyzed quantitatively and qualitatively by real-time PCR (qRT) and overlapping reverse transcription (RT) PCRs, respectively. *Results:* All VWF tests were normal. The LR-PCRs demonstrated the integrity of the *F8* locus. Eight unique polymorphisms were found in the patients, with two being recurrent. Furthermore, RT-PCRs analysis confirmed that two of the unique variants create detectable new cryptic splice sites in the patients that result in the introduction of intronic DNA sequences into the mRNA and create premature stop codons. *Conclusion:* By systematically excluding all possible causes of HA, we could with great certainty conclude that deep intronic mutations in *F8*, although rare, cause abnormal mRNA splicing, leading to mild HA.

Keywords: factor VIII; hemophilia A; introns; mRNA; mutations.

Correspondence: Osman El-Maarri, Institute of Experimental Hematology and Transfusion Medicine, University of Bonn, Sigmund-Freud-Str. 25, 53125, Bonn, Germany.
Tel.: +49 228 28716737; fax: +49 228 28714320.
E-mail: Osman.Elmaarri@ukb.uni-bonn.de

Received 26 March 2013
Manuscript handled by: L. Aledort
Final decision: F. R. Rosendaal, 23 June 2013

Introduction

HA is a monogenetic disease caused by a broad spectrum of mutations in the factor (F) VIII gene (*F8*) (HAMSTeRS; www.hadb.org.uk). Most known genetic lesions leading to HA are within the *F8* cDNA. Moreover, mutations in certain domains of von Willebrand factor (VWF) that interact with FVIII cause an HA-like phenotype known as 2N VWD [1]. However, in about 2% of patients with reduced FVIII:C levels presenting a typical moderate to mild HA phenotype, no mutations could be identified in *F8* cDNA [2].

Since the discovery of the intronic regions, there has been an increase in understanding splicing errors and their role in pathology [3]. Traditionally, mutations affecting the consensus splice-sites have been identified using a routine intron/exon boundary screening. Once regarded as neutral or even dispensable, intronic sequences are now known to dictate the correct splicing of the pre-mRNA through several elements, such as branch point sequences, polypyrimidine tract and splice site consensus elements [4]. Therefore, mutations or rearrangements in the non-coding DNA have to be considered as a potential cause of the HA phenotype in patients without mutations. Indeed, several other studies reported the presence of causal intronic mutations. Such intronic mutations result in an altered mRNA: either by exclusion of one or several exons causing a severe HA phenotype [5] or through intron inclusion leading to a mild HA phenotype [6–9]. The first is caused by mutations located within or close to

intron/exon boundaries, whereas the latter are found in deep intronic regions that are not detected while applying the routine mutation screening methods [6,8].

The aim of this study was to establish a universal systematic approach to investigate the molecular mechanism causing reduced FVIII:C levels in a group of mild (only one moderate) HA patients without a mutation in *F8* cDNA. As a result we established a four-step protocol. First step: amplification of the whole F8 locus using overlapping LR-PCRs. Second step: NGS of LR-PCR fragments and identification of patient's unique variants. Third step: qualitative RT-PCR to detect abnormal splicing. Fourth step: verifying the reduction of F8 mRNA at a given exon-exon junction flanking the intron comprising the unique variant by qRT-PCR. Using this protocol the causative intronic mutation was identified in two patients.

Materials and methods

Samples

A total of 96 non-hemophilic males and nine HA patients with no apparent genomic mutations in *F8* cDNA are included in this study; in addition, a further 96 healthy males and 96 healthy females were used for screening of the rare SNPs. Patients included in this study were previously tested for intron 1 and 22 inversions and the recently reported intron 1 (Int1R) rearrangement [10–12]. The study was approved by the ethics committee of the University of Bonn (approval number 183/07).

DNA isolation and mutation analysis

Genomic DNA was extracted from peripheral blood collected from HA patients and healthy controls. Mutation analysis was performed for all coding exons of *F8*, *VWF*, *LMAN1* and *MCFD2*. For duplication detection, MLPA was carried out using the F8-P178 kit (MRC-Holland, Amsterdam, the Netherlands).

F8 mRNA analysis

RNA isolation from peripheral blood and amplification of *F8* mRNA were carried out as described earlier [2]. For qRT-PCRs, a TaqMan-based assay was performed using the AgPath-ID Kit (Life Technologies, Darmstadt, Germany) according to the manufacturer's instructions. *PBGD* (NM_000190) and the gene expression assay Hs00240767_m1 (Life Technologies) were used as internal controls to normalize *F8* mRNA expression (Table S1). The allelic analysis of mRNA was carried out as described earlier [13].

Next generation sequencing (NGS)

The complete *F8* locus was amplified in 28 overlapping PCR regions as described previously [10]. NGS was performed on an Illumina Genome Analyzer II using standard protocols (Illumina Inc., San Diego, CA, USA) (Fig. 1). Briefly, for each patient the LR-PCR fragments were equimolarly mixed and processed after fragmentation according to Illumina's Paired-End Sequencing library kits. Prior to that, each patient's library was tagged with a unique barcode. The Paired-End adapter-ligated fragments of the pooled libraries (200 bp $\pm$ 50 bp) were attached to the flow cell and amplified by bridge amplification via PCR. The generated clusters were sequenced and reads were aligned onto the reference assembly NG_011403.1, allowing at most 4 mismatches/36 bases using Geneious software (Biomatters Ltd, Auckland, New Zealand). The coverage varied between 72 and 692-fold. A minimum of 94.6% of the *F8* locus was covered after NGS. After analysis a list of variants was generated and the significance of base calling was determined by sorting out false-positive calls by verification of five variants with a frequency call above 50%. A threshold of 90% call was found to be the false-positive border and a minimum coverage of three reads was the cut-off corresponding to a real variation.

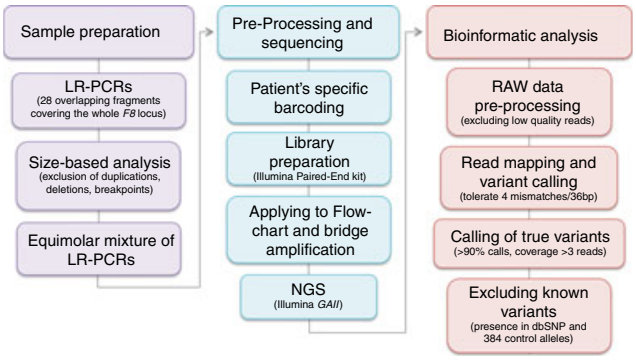

Fig. 1. Flowchart of the NGS approach.

Splice site prediction analysis

Wild-type (wt) and mutant sequences were submitted for splice site prediction to the Berkeley Drosophila Genome Project (BDGP) (http://fruitfly.org/seqtools/splice.html), NetGene2 server (http://www.dtu.dk/services/NetGene2/) and Human Splicing Finder server (http://www.umd.be/HSF/).

X-chromosome inactivation (XIR) and likelihood of carriership

The XIR test was performed using the *HUMARA* assay as described by Allen *et al*. [14]. The likelihood of being a carrier was determined using FVIII:C values after correction for the errors/deviations introduced by age and blood group. The transformed value ($T_{FVIII:C}$) was calculated according to the formula $T_{FVIII:C} = \ln(\text{FVIII:C})\text{-age}^2 \times 0.00011\text{-bg} \times 0.18017\text{–}3.61700$ as described by Oldenburg *et al*. [15], where bg = 0 if the subject has blood group O or A2 and bg = 1 if A1, B or AB. The *T* values were used to extrapolate the likelihood of being a carrier or non-carrier.

Results

The clinical characteristics of typical HA were reconfirmed in all patients (Table 1). Mutations in *LMAN1*, *MCFD2* and *VWF* were excluded (Table S2). Functional defects of VWF protein were excluded by performing VWF:CB, VWF:FVIIIB and multimer analysis (Table 1; Figure S1). MLPA analysis revealed no duplications/deletions in *F8*. Finally, translocation of *F8* was excluded via FISH analysis using a *F8*-specific probe (covering exons 2-26; BAC number IMGSB737DO62145D from ImaGenes, Berlin, Germany) in three patients (50P, 79P and 88P) (data not shown).

Identification of unique deep intronic variations in HA patients

Forty-four intronic variations were found to fulfill the detection threshold; 36 of them were reported in the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) and/or were present in our initial cohort of 96 healthy male controls (96 alleles) (Table S3B). The remaining eight unique intronic variations (Table S3C) were absent in an additionally screened 288 alleles (96 male and 96 female controls) (data not shown). Interestingly, two of the unique variants were recurrent in the analyzed patients; SNP2 (c.2113+11641insG) was present in 3P and 15P, and SNP3 (c.5586+194C>T) was present in 19P, 77P and 88P.

All eight unique variations were predicted to create new donor and/or acceptor splice sites (dss/ass) by at least one of the prediction softwares, except for SNP1 (c.2113+8034C>G) and SNP8 (c.6429+28519A>G). For two SNPs [SNP5 (c.5998+941C>T, 83P) and SNP6 (c.6273+2864A>C, 79P)], the generation of new ass and dss were predicted in intron 18 and intron 21, respectively. However, the predicted strength of the new splice site, according HSF, was lower than that of the wt site (Table 2). For all other SNPs, the new ass and dss scores were higher than the wt sequence. The inconsistency between different algorithms in predicting the effect of specific variations on splicing necessitates the verification by experimental approaches.

RT-PCR analysis reveals aberrant splicing for two intronic variations

In two patients (79P and 83P), aberrant mRNA transcripts were observed in region C2 (exons 17-21) (Fig. 2 and S4). Sequencing of these products revealed the inclusion of intronic sequences of intron 18 (105 bp and 126 bp insertion for 79P and 83P, respectively). Analysis of the flanking sequences of these insertions at the DNA level revealed the presence of the two unique intronic variations identified by NGS (c.5998+530C>T and c.5998+941G>A for 79P and 83P, respectively). The new dss for the intronic mutation in 79P was highly predicted by two of the prediction algorithms. However, for 83P, the intronic variant appears to create a false splice donor site in intron 18 of *F8*; AG>AA at +950 bases from exon 18. None of the splice site prediction programs predicted this newly created site. Our sequencing results showed that this transition activates another splice site +10 bp downstream of the mutation.

Quantitative mRNA analysis reveals the reduction of wt F8 mRNA across exons flanking intronic mutations

As we could not find aberrant mRNA splicing in all patients with unique intronic variations, we performed real-time PCR analysis to measure the amount of *F8* mRNA at exon boundaries where an alternative splicing was suspected. The assay was designed for a control region (exon 1-2 of *F8*) located upstream of all unique SNPs and a test region for exon boundaries flanking an intronic variation. The ratio of *F8* mRNA expression at the control region to that at the tested region was calculated and compared between the patients and healthy controls. A reduced ratio in comparison to the controls indicates the presence of alternative splicing at the investigated region. To validate this approach, the two intronic mutations that were proven to result in abnormal spliced products were used (79P and 83P, Fig. 2D). For both variations, a significant reduction in the amount of *F8* mRNA in the test region was observed (Mann-Whitney test, $P = 0.0005$ for c.5998+530C>T and $P = 0.0004$ for c.5998+941G>A) (Fig. 2D). This result is in accordance with the presence of the wild-type mRNA in both patients as a faint band in the RT-PCRs (Fig. 2C: red arrows).

Table 1 Summary of laboratory measurements and clinical profile of studied patients

	Patients			3P	15P	50P	84P	83P	19P	77P	79P	88P
	Clinical severity			Mild	Mild	Mod.	Mild	Mild	Mild	Mild	Mild	Mild
			Normal range									
Coagulation parameters	FVIII	Activity	57–245%	29	21	3	5.3	9.9	14	35	7.4	12.3
		Antigen	50–183%	–	34	4	7.3	14.5	–	9.9	22.1	18.8
		Sp. Act. *	0.67–2.05	–	0.62	0.75	0.73	0.68	–	3.53	0.34	0.65
		Expression†	0.028–0.75	0.122	0.263	0.065	0.062	0.038	0.038	0.056	0.046	0.033
	VWF	Activity	44–276%	86	53	76	80	168	110	60	132	80
		Antigen	40–250%	79	70	90	108	164	71	59	170	89
		Sp. Act.	0.406–1.844	1.09	0.76	0.84	0.74	1.02	1.54	1.02	0.78	0.9
		Collagen binding	50–160%	–	63	61	87	97	–	46	118	87
		FVIII binding	60–170%	–	103	95	96	100	–	98	126	106
	FII	Activity	87–127%	–	101	105	97	88	–	–	95	103
	FV	Activity	67–145%	–	134	86	122	119	–	148	111	123
	FVII	Activity	77–154%	–	119	110	115	137	–	–	88	55
	FX	Activity	80–140%	–	108	133	125	157	–	–	92	–
	Anti-thrombin	Activity	85–130%	–	99	98	96	92	–	–	99	87
		aPTT	25–36 sec.	–	34.3	46.6	43.6	42.9	–	–	44.5	32.6
	FVIII-inhibitor‡			Neg.	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.
	DDAVP response after 30s§			–	4.17	–	0.98	–	–	–	–	6.5
Clinical parameters	Blood group			A2	O	A	B	B	O	O	A1	O
	Medication	Rec. or Pd.		–	Rec.	Rec.	Rec.	Rec.	–	–	Pd.	Rec.
		VWF presence		–	No	No	No	No	–	–	Yes	No
	FVIII:C	24 h	Amount IU	–	3000	2000	3000	3000	–	–	3000	3000
			FVIII:C	–	24%	29%	17%	21%	–	–	40%	26%
		48 h	Amount IU	–	2000	–	–	2000	–	–	–	3000
			FVIII:C	–	16.5%	–	–	13.5%	–	–	–	18.5%
	Bleeding tendency¶			–	N	N	N	N	–	–	N	N
	Amount of factors used**			–	↓	↓	↓	↓	–	–	↑	↓

Neg., negative; IU, International units; Rec., recombinant FVIII-; Pd., plasma-derived concentrates; ↓, below average; ↑, above average; N, adequate. *Specific activity is calculated as the ratio of activity to antigen. †Expression of *F8* mRNA by qRT-PCR. ‡Based on Bethesda assay Neg.=<0.6I U/ml. §Response to DDAVP: mild, 0.8–16; moderate, 0.8–156. ¶General evaluation of bleeding tendency. **Amount of therapeutically infused factors in comparison with patients with known mutations.

Table 2 Analysis of unique SNPs found in patients without mutation in *F8* cDNA

Patients	3P		15P		83P		19P	77P		79P			88P
SNP													
Intron	13	22	13	13		18	16	16	18	21	22		16
Variant	c.2113	c.6429	c.2113	c.2113	c.5998		c.5586	c.5586	c.5998	c.6273	c.6429	c.5586	
	+11641insG	+28519G>T	+11641insG	+8034C>G	+941G>A		+194C>T	+194C>T	+530C>T	+2864A>C	+6461G>T	+194C>T	
SNP Nr. *	2	8	2	1	5		3	3	4	6	7		3
Splicing prediction													
FRUITFLY (0-1)	np	np	np	np	np		D: mut = 0.88	D: mut = 0.89	D: mut = 0.99	np	np		D: mut = 0.9
NETGENE2 (0-1)	np	np	np	A: mut = 0.15	np		D: mut = 0.56	D: mut = 0.57	D: mut = 0.35	D: mut = 0.45	np		D: mut = 0.58
HSF (0-100)	D:	np	D:	Np	A:		A:	A:	D:	D:	D:		A:
	44.8 > 72.71		44.8 > 72.71		75.72 > 46.77		65.56 > 66.41	65.56 > 66.42	59.3 > 86.13	65.68 > 60.93	51.3 > 80.25	65.56 > 66.43	
	A:	np	A:	np	np		D:	D:	np	np	np	D:	
	19.1 > 72.76		19.1 > 72.76				63.06 > 89.9	63.06 > 89.10				63.06 > 89.11	
RNA analysis													
RT-PCR	No RNA	No RNA	Normal	Normal	Insertion		Normal	Normal	Insertion	Normal	Normal	Normal	
qRT	No RNA	No RNA	Reduced	Normal	Reduced		ND	Reduced	Reduced	Normal	Normal	ND	

D, dss; A, ass; RT-PCR, qualitative mRNA analysis; qRT, exon-exon specific quantitative TaqMan-based RT-PCRs; ND, not determined; np, not predicted; mut, mutant. *Table S3C.

Moreover, a reduction of *F8* mRNA expression (Mann–Whitney test, $P = 0.019$) for patient 77P was observed in the region spanning the exons flanking SNP3 (c.5586+194C>T; also present in 19P and 88P, Figure S5A). These data suggest the presence of an aberrant mRNA in this patient as well. However, we could not detect this using our nested RT-PCR approach. Probably, this alternative splicing involves larger regions beyond the current amplification limits of the RT-PCR approach.

For SNP2 (c.2113 + 11641insG in intron 13) present in 15P (and 3P), no difference in the amount of *F8* mRNA was observed when analyzing the exon 13-14 junction. This result suggests that if this SNP is causal then through a still elusive mechanism.

Association study of non-unique SNPs with FVIII:C, FVIII:Ag and F8 expression

Furthermore, the effect of extra- [16] and intra-genetic polymorphisms (detected after NGS) on reducing FVIII:C was analyzed. Only one SNP showed borderline significance (rs1800292 [c.3864A>C, $P = 0.0627$]). Moreover, the effect of all 36 non-unique polymorphisms on *F8* mRNA expression, FVIII:Ag and FVIII:C was analyzed by performing association studies (Table S3 and Figure S2). A significant association was found between *F8* mRNA expression and rs28857481 and rs6643622 ($P = 0.0326$ and $P = 0.0071$, respectively). However, no significant association with either FVIII:Ag nor FVIII:C was found for these two SNPs. Among all SNPs, only one (rs34552198, c.2114-7226_2114-7225insTA) was significantly associated with FVIII:C ($P = 0.0490$), while two showed borderline significance (rs78327897 (c.6430-3402G>A), $P = 0.0611$ and rs1050705 (c*1672G>A), $P = 0.0535$).

All four SNPs that showed significant or borderline significant association with FVIII:C, were used to construct haplotypes. In total, six haplotypes were found (Figure S3). When stratified for the ABO polymorphism, the AA and AG genotypes showed significant differences in FVIII:C levels among different haplotypes ($P = 0.0048$). The H4 and H5 haplotypes showed significant association with reduced FVIII:C levels in comparison to H1, H2 and H3 haplotypes. However, all studied patients have haplotypes associated with higher FVIII:C (3P and 15P have haplotype H3, while the rest of the patients have haplotype H1). This excludes the effect of common polymorphisms in *F8* on reducing FVIII:C in the studied patients.

For two available pedigrees index patients and potential carriers share the same F8 allele

Family 1 The index patient shares the same allele with five potential carriers (I-2, II-4, II-5, III-1 and III-6), only one of whom shows significant reduction in FVIII:C

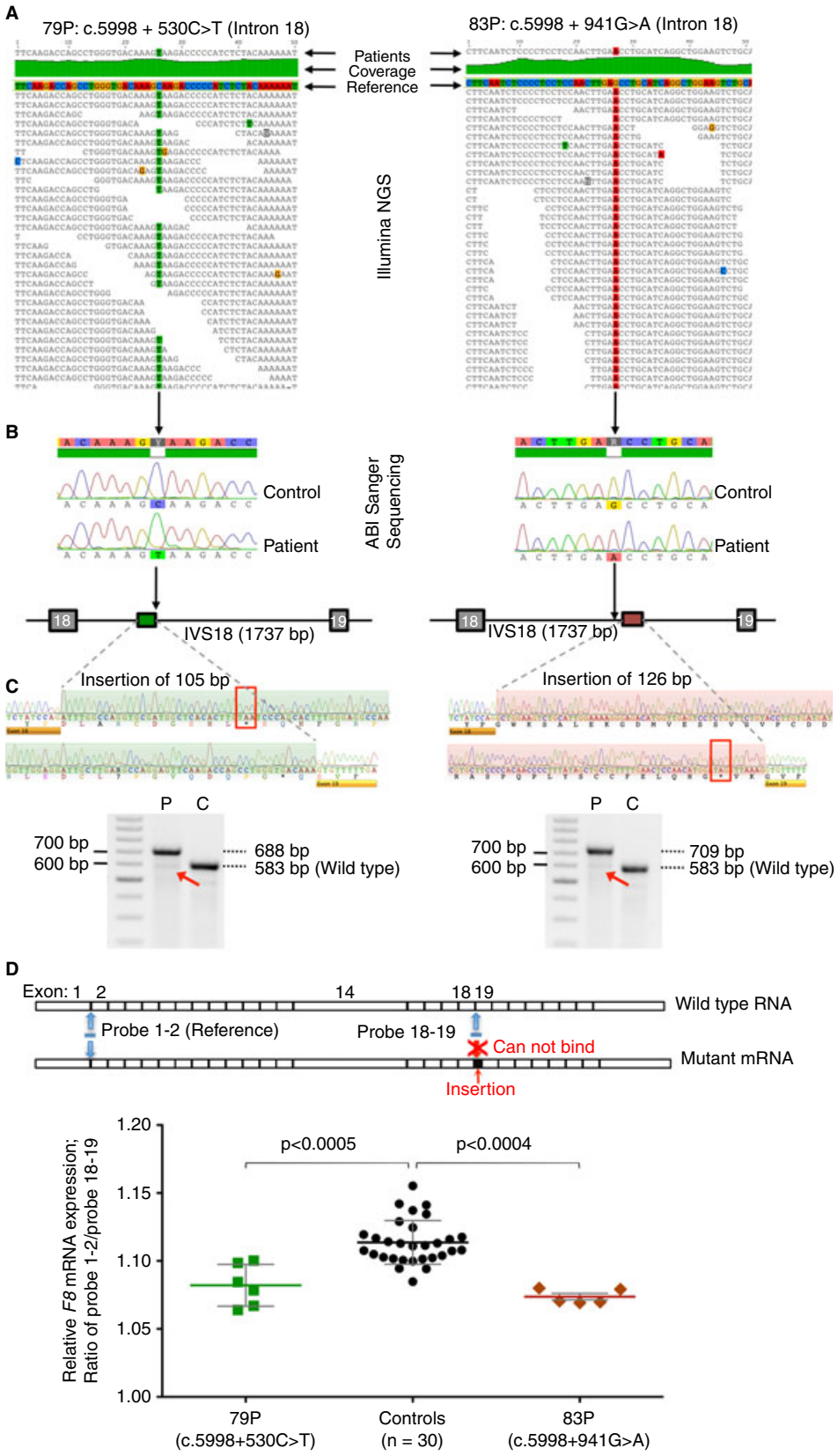

Fig. 2. Identification and characterization of intronic mutations for 79P and 83P. (A) Screenshot of NGS data. (B) Verification of the mutation by Sanger sequencing (upper panel) and schematic presentation of relative position of the intronic mutations (lower panel). (C) The inserted sequences caused by the mutations (upper panel) with RT-PCR pictures (lower panel). The red arrows indicate the faint wt mRNA band. (D) qRT-PCR results with the positions of TaqMan probes (upper panel) and the ratio of control region to the tested region (lower panel).

(III-6: 35%) (Figure S6A). For all potential carriers, the XIR pattern was determined to explain the reason for differences in FVIII:C levels despite carrying the allele of the index patient. The *HUMARA* assay in potential carriers with normal FVIII:C levels (I-2, II-4, II-5 and III-1) showed inactivation of the index patient's allele in 82%, 100%, 95% and 69% of the cells, respectively, explaining their relative normal FVIII:C values. Interestingly, for III-6, about 50% of cells were expressing the patient's allele, explaining the reduction in FVIII:C. Based on an informative marker in exon 14 of *F8*, allelic analysis was performed for III-1 (daughter of the index patient). Using this assay we observed that only the non-index patient's allele was detected on RNA level. This result supports the normal FVIII:C in accordance with the XIR pattern (Figure S7).

Family 2 This family includes two affected members and three obligate carriers (Figure S6B). The index patient (84P) has a mild phenotype with FVIII:C of 5.3% and FVIII:Ag of 7.3%, while the other affected member (50P) has a moderate phenotype with FVIII:C and FVIII:Ag of 2.6% and 4%, respectively. Based on haplotype analysis all patients and obligate carriers share the same *F8* allele. Moreover, all potential carriers show a reduction in FVIII:C in accordance with their XIR pattern.

Discussion

In a small group of clinically defined non-severe HA patients, even after applying various mutation detection techniques, no mutations could be detected in *F8* cDNA [2,9,17]. A number of mutations are thus being missed, as some of these will involve sequence variations in intronic regions that are not being assessed by routine diagnosis. The mRNA analysis performed in this study combined with NGS is a convenient complementary method to prove the disease association of intronic mutations.

It is becoming more evident that deep intronic sequence variations affect the splicing process. Deep intronic sequence 'exonisation' has been reported for several genes, such as *CFTR* [18], *BRCA2* [19] and *HRPT* [20]. Also for the above genes, the importance of screening for splicing variations has been underlined in *F8*. Castaman *et al*. presented three different intronic variations (c.2113+1152delA, c.5587-93C>T and c.5999-277G>A) resulting in generation of new splice sites and premature stop codons [6]. A single transition within intron 10 (c.1537+325A>G) was reported by Inaba *et al*. as a causal mutation in a patient with mild HA [8]. Such intronic mutations activate a cryptic donor or acceptor splice site. Splicing between this novel splice site and a pre-existing splice site leads to the inclusion of a cryptic exon or simply to alternative splicing.

Our mRNA analysis includes two approaches: qualitative RT-PCRs to detect aberrant splicing and qRT-PCR to determine the relative amount of *F8* mRNA expression at an exon-exon junction flanking an intronic variation. It should be taken into consideration that as the cDNA used in this study is synthesized from RNA obtained from blood, which is not a known production site of FVIII protein, possible tissue-specific splicing cannot be neglected [21]. Therefore there is no assurance that the effect of the variations found here will be the same in FVIII expressing/secreting cells. Nevertheless, the importance of screening for splicing variations has been underlined in HA [2,5,6,13] and our work highlights the importance of implementation of *F8* mRNA analysis in clinical practice.

Using our combined NGS and mRNA analysis protocol, we could determine the exact molecular defect leading to mild HA in two patients (83P and 79P). In both cases, intronic variations created new cryptic sites that led to the insertion of intronic sequences in *F8* mRNA between exon 18 and 19. Both insertions introduce premature stop codons in *F8* mRNA that produce dysfunctional FVIII protein. These truncated proteins contribute to the slightly higher ratio of FVIII:Ag to FVIII:C levels in both patients. Moreover, the mild phenotype is well explained by the presence of a relatively low amount of normally spliced *F8* mRNA (demonstrated by RT-PCRs).

In a group of three patients (19P, 77P and 88P) a recurrent unique variant in intron 16 (c.5586+194C>T) was found. This variation did not lead to detectable mRNA alternative splicing but to a significant reduction of *F8* mRNA expression across exons 16-17. This could indirectly indicate that there is an abnormal mRNA across this intron, which makes this variant highly likely to be causative for three reasons: (i) absence in the healthy population, (ii) it is recurrent in three unrelated patients and (iii) the level of wt *F8* mRNA is reduced across the intron harboring this variant. However, there is a possibility of the existence of a defect in still not identified regulatory regions or an impaired transcription of FVIII at its main site of production, which is the liver.

In two patients (3P and 15P), a second recurrent unique variant was detected (c.2113+11641insG). Both qRT and qualitative mRNA analysis did not reveal abnormalities. Although, on one hand its recurrence in the patient population and absence in the healthy population is suggestive of its causality, on the other hand the absence of a detectable effect on *F8* mRNA is suggestive of a benign effect. Therefore, the causality of this variation could not be proven. The link between the variant and the mutation mechanism could be more complex.

In the two remaining patients (50P, 84P), we could not detect any DNA change or rearrangement across the *F8* locus including the introns. However, haplotype analysis revealed that all affected patients and potential carriers share the same *F8* allele. This result is suggestive of a defect in *F8*. However, the nature of this defect is still elusive.

Taken together, the study presented here highlights the necessity of looking beyond the coding, intron/exon boundary and promoter regions, when attempting to establish the molecular cause of a disease in the absence of exonic mutations. We show that LR-PCRs and NGS are likely to be useful in a diagnostic setting to clarify the genetic basis of patients in whom no mutation through conventional Sanger sequencing is detectable. We show that deep intronic mutations in *F8*, although rare, can cause abnormal splicing that leads to HA. We present a comprehensive four-step protocol to investigate the causality of such intronic mutations. Our approach is universal and could be applied to any monogenic disease.

Addendum

O. El-Maarri and B. Pezeshkpoor designed the study. B. Pezeshkpoor, N. Zimmer and I. Nanda performed the experiments. B. Pezeshkpoor, N. Marquardt, T. Haaf, U. Budde, J. Oldenburg and O. El-Maarri analyzed and interpreted data. B. Pezeshkpoor and O. El-Maarri wrote the manuscript.

Acknowledgements

This work was supported by a DFG grant (Deutsche Forschungsgemeinschaft: EL499/2-1), a Baxter bioscience grant (Number: H12-000820) and the Bayer Haemophilia award program.

Disclosure of Conflict of Interests

The authors state that they have no conflict of interests to declare.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Normal VWF multimer pattern was observed in the analyzed patients. Samples were diluted according to their VWF content to approximately 0.5 μg/ml.
Figure S2. Association between SNPs and *F8* mRNA expression, FVIII:Ag and FVIII:C in the healthy population.
Figure S3. Haplotype analysis based on four SNPs associated with FVIII:C. Activity levels are plotted for all six haplotypes. (A) non-stratified and (B) stratified for the rs687289.
Figure S4. Schematic diagrams representing the amplified RT-PCR fragments (upper part) with their corresponding experimental gel pictures of patients (lower part).
Figure S5. qRT-PCR analysis for two unique recurrent intronic variants. (A) c.5586+194C>T and (B) c.2113+11641insG.
Figure S6. Haplotype analysis in pedigrees of (A) family 1 and (B) family 2.
Figure S7. HPLC chromatograms of the single nucleotide primer extension product of the allelic analysis at the DNA and RNA level at codon 1269 (rs1800292, c.3864C>A) for 15P and his daughter (Figure S6, Family 1).
Table S1. List of primers used in this study.
Table S2. Detected polymorphisms in *VWF*.
Table S3. List of SNPs analyzed for correlation with *F8* mRNA expression and FVIII and VWF activity and antigen levels in healthy controls and patients.

References

1 Mazurier C, Dieval J, Jorieux S, Delobel J, Goudemand M. A new von Willebrand factor (vWF) defect in a patient with factor VIII (FVIII) deficiency but with normal levels and multimeric patterns of both plasma and platelet vWF. Characterization of abnormal vWF/FVIII interaction. *Blood* 1990; **75**: 20–6.
2 El-Maarri O, Herbiniaux U, Graw J, Schroder J, Terzic A, Watzka M, Brackmann HH, Schramm W, Hanfland P, Schwaab R, Muller CR, Oldenburg J. Analysis of mRNA in hemophilia A patients with undetectable mutations reveals normal splicing in the factor VIII gene. *J Thromb Haemost* 2005; **3**: 332–9.
3 Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* 2007; **8**: 749–61.
4 Zhang N, Kaur R, Lu X, Shen X, Li L, Legerski RJ. The Pso4 mRNA splicing and DNA repair complex interacts with WRN for processing of DNA interstrand cross-links. *J Biol Chem* 2005; **280**: 40559–67.
5 Castaman G, Giacomelli SH, Mancuso ME, Sanna S, Santagostino E, Rodeghiero F. F8 mRNA studies in haemophilia A patients with different splice site mutations. *Haemophilia* 2010; **16**: 786–90.
6 Castaman G, Giacomelli SH, Mancuso ME, D'Andrea G, Santacroce R, Sanna S, Santagostino E, Mannucci PM, Goodeve A, Rodeghiero F. Deep intronic variations may cause mild hemophilia A. *J Thromb Haemost* 2011; **9**: 1541–8.
7 Santacroce R, Santoro R, Sessa F, Iannaccaro P, Sarno M, Longo V, Gallone A, Vecchione G, Muleo G, Margaglione M. Screening of mutations of hemophilia A in 40 Italian patients: a novel G-to-A mutation in intron 10 of the F8 gene as a putative cause of mild hemophilia A in southern Italy. *Blood Coagul Fibrinolysis* 2008; **19**: 197–202.
8 Inaba H, Koyama T, Shinozawa K, Amano K, Fukutake K. Identification and characterization of an adenine to guanine transition within intron 10 of the factor VIII gene as a causative mutation in a patient with mild haemophilia A. *Haemophilia* 2013; **19**: 100–5.
9 Hallden C, Knobe KE, Sjorin E, Nilsson D, Ljung R. Investigation of disease-associated factors in haemophilia A patients without detectable mutations. *Haemophilia* 2012; **18**: e132–7.
10 Pezeshkpoor B, Rost S, Oldenburg J, El-Maarri O. Identification of a third rearrangement at Xq28 that causes severe hemophilia A as a result of homologous recombination between inverted repeats. *J Thromb Haemost* 2012; **10**: 1600–8.
11 Liu Q, Nozari G, Sommer SS. Single-tube polymerase chain reaction for rapid diagnosis of the inversion hotspot of mutation in hemophilia A. *Blood* 1998; **92**: 1458–9.
12 Bagnall RD, Waseem N, Green PM, Giannelli F. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood* 2002; **99**: 168–74.
13 El-Maarri O, Singer H, Klein C, Watzka M, Herbiniaux U, Brackmann HH, Schroder J, Graw J, Muller CR, Schramm W,

Schwaab R, Haaf T, Hanfland P, Oldenburg J. Lack of F8 mRNA: a novel mechanism leading to hemophilia A. *Blood* 2006; **107**: 2759–65.

14 Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet* 1992; **51**: 1229–39.

15 Oldenburg J, Schwaab R, Grimm T, Zerres K, Hakenberg P, Brackmann HH, Olek K. Direct and indirect estimation of the sex ratio of mutation frequencies in hemophilia A. *Am J Hum Genet* 1993; **53**: 1229–38.

16 Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, Hayward C, Rudan I, Sabater-Lleal M, Bis JC, de Maat MP, Rumley A, Kong X, Yang Q, Williams FM, Vitart V, Campbell H, Malarstig A, Wiggins KL, van Duijn CM, *et al*. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: the CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation* 2010; **121**: 1382–92.

17 Klopp N, Oldenburg J, Uen C, Schneppenheim R, Graw J. 11 hemophilia A patients without mutations in the factor VIII encoding gene. *Thromb Haemost* 2002; **88**: 357–60.

18 Chillon M, Dork T, Casals T, Gimenez J, Fonknechten N, Will K, Ramos D, Nunes V, Estivill X. A novel donor splice site in intron 11 of the CFTR gene, created by mutation 1811+1.6kbA–>G, produces a new exon: high frequency in Spanish cystic fibrosis chromosomes and association with severe phenotype. *Am J Hum Genet* 1995; **56**: 623–9.

19 Anczukow O, Buisson M, Leone M, Coutanson C, Lasset C, Calender A, Sinilnikova OM, Mazoyer S. BRCA2 deep intronic mutation causing activation of a cryptic exon: opening toward a new preventive therapeutic strategy. *Clin Cancer Res* 2012; **18**: 4903–9.

20 Corrigan A, Arenas M, Escuredo E, Fairbanks L, Marinaki A. HPRT deficiency: identification of twenty-four novel variants including an unusual deep intronic mutation. *Nucleosides, Nucleotides Nucleic Acids* 2011; **30**: 1260–5.

21 Xu Q, Modrek B, Lee C. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res* 2002; **30**: 3754–66.