

RESEARCH ARTICLE

Pathological mechanism and antisense oligonucleotide-mediated rescue of a non-coding variant suppressing factor 9 RNA biogenesis leading to hemophilia B

Simon Krooss^{1,2}, Sonja Werwitzke³, Johannes Kopp¹, Alice Rovai², Dirk Varnholt³, Amelie S. Wachs¹, Aurelie Goyenvalle⁴, Annemieke Aarstma-Rus⁵, Michael Ott², Andreas Tiede³, Jörg Langemeier^{1,6*}, Jens Bohne^{1*}

1 Institute of Virology, Hannover Medical School, Hannover, Germany, **2** Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School and Twincore Centre for Experimental and Clinical Infection Research, Hannover, Germany, **3** Clinic of Hematology, Oncology and Hemostaseology, Hannover Medical School, Hannover, Germany, **4** Université de Versailles St-Quentin en Yvelines, INSERM U1179, France, **5** Leiden University Medical Center, Leiden, Netherlands, **6** Pediatric Intensive Care Unit, Children's Hospital Bielefeld, Germany

* joerglangemeier@web.de (JL); bohne.jens@mh-hannover.de (JB)



OPEN ACCESS

Citation: Krooss S, Werwitzke S, Kopp J, Rovai A, Varnholt D, Wachs AS, et al. (2020) Pathological mechanism and antisense oligonucleotide-mediated rescue of a non-coding variant suppressing factor 9 RNA biogenesis leading to hemophilia B. *PLoS Genet* 16(4): e1008690. <https://doi.org/10.1371/journal.pgen.1008690>

Editor: Gregory M. Cooper, HudsonAlpha Institute for Biotechnology, UNITED STATES

Received: September 10, 2019

Accepted: February 22, 2020

Published: April 8, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1008690>

Copyright: © 2020 Krooss et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Abstract

Loss-of-function mutations in the human coagulation factor 9 (*F9*) gene lead to hemophilia B. Here, we dissected the consequences and the pathomechanism of a non-coding mutation (c.2545A>G) in the *F9* 3' untranslated region. Using wild type and mutant factor IX (FIX) minigenes we revealed that the mutation leads to reduced *F9* mRNA and FIX protein levels and to lower coagulation activity of cell culture supernatants. The phenotype could not be compensated by increased transcription. The pathomechanism comprises the *de novo* creation of a binding site for the spliceosomal component U1snRNP, which is able to suppress the nearby *F9* poly(A) site. This second, splicing-independent function of U1snRNP was discovered previously and blockade of U1snRNP restored mutant *F9* mRNA expression. In addition, we explored the vice versa approach and masked the mutation by antisense oligonucleotides resulting in significantly increased *F9* mRNA expression and coagulation activity. This treatment may transform the moderate/severe hemophilia B into a mild or subclinical form in the patients. This antisense based strategy is applicable to other mutations in untranslated regions creating deleterious binding sites for cellular proteins.

Author summary

The elucidation of the pathomechanisms of non-coding variants yields important insights into diseases as well as cellular processes causing the defect. Although these variants may account for the majority of phenotypic variation, only a minority of them can be explained mechanistically. The human coagulation factor 9 3' UTR variant described here converts a non-essential sequence motif into a U1snRNP-binding site with deleterious effects on

Funding: This work was supported by the Else-Kröner-Fresenius-Stiftung (2014_A88) to J. B. and the German research foundation (DFG BO2512/6-1) to J. B. and (DFG OT131/6-1) to M. O. S. K. was supported by the Studienstiftung des Deutschen Volkes. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

RNA 3' end processing at the nearby poly(A) site. Poly(A) site suppression by U1snRNP was described before and it normally protects cellular mRNAs from premature termination. However, if misled by creation of a U1 site close the authentic poly(A) site as in the *F9* 3' UTR, this nuclear surveillance mechanism results in the opposite. Since recognition by U1snRNP depends on sequence complementarity we were able to use antisense oligonucleotides to mask the mutant site and partially restored *F9* mRNA levels. This antisense based strategy may be applicable to other variants in untranslated regions, which create deleterious binding sites for cellular proteins.

Introduction

Variants in the non-coding part of the genome can result in severe diseases and phenotypes with a high expressivity [1] and the number of variants identifiable by GWAS can explain only a modest fraction of the phenotypic variation [2]. However, the pathomechanism underlying these variants are often difficult to dissect.

Factor IX (FIX) constitutes a serine protease in the enzyme cascade of the intrinsic coagulation pathway. The *F9* gene is located on the X chromosome consisting of 8 exons [3]. Disease-causing variants in the *F9* gene cause complete or partial deficiency of the corresponding coagulation factor, ultimately leading to hemophilia B (OMIM:306900) of varying severity with an incidence of 1 in 25.000 births [4]. The FIX mutation database (<http://www.factorix.org>); [5] lists 1083 different variants. Of these, 122 are non-coding and four occur in the 3' untranslated region (3'UTR). These non-coding variants can have severe consequences and in addition, they are often part of regulatory sequences, which otherwise would escape our notice [1].

In 1993 a 3'UTR variant (Chr.X, NC_000023.11:g.139563228A>G) was described in four families with moderate to severe hemophilia B [6]. Affected individuals showed FIX activity between <1 and 4.5% of normal. At present, the FIX mutation database lists 21 patients from different geographic regions with this variant (<http://www.factorix.org>); [7, 8]; S1 Table). The variant may generate a 5' splice site (5'SS) by an A to G transition at position +1157 downstream of the stop codon (c.2545A>G; NM_000133) in the 3' UTR [6]. However, the underlying pathophysiological mechanism remained unsolved.

Eukaryotic pre-mRNAs undergo multiple processing steps. The removal of introns is executed by the spliceosome, which consists of five small nuclear U-rich ribonucleoproteins (UsnRNPs). The 5'SS is recognized by the RNA component of the U1snRNP [9]. The strength of a U1 binding site is primarily determined by its sequence complementarity to U1snRNA [10]. In addition to splicing, mRNAs are cleaved and polyadenylated at their 3' end. Again, several conserved sequence features in close proximity to the hexameric poly(A) signal (PAS) direct the assembly of the cleavage and polyadenylation (CPA) machinery. The CPA is confronted with the task to identify the correct PAS among many close matches. Especially the long mammalian introns harbor many cryptic PAS. Recently, we and others have shown that U1snRNP is also part of a nuclear surveillance mechanism that suppresses cryptic polyadenylation sites and thereby ensures transcript integrity [11, 12].

For the *LAMTOR2* (p14) gene, we demonstrated that a 3'UTR variant creating a U1 binding site leads to PAS suppression, thereby causing an immunodeficiency [12]. Thus, the generation of a *de novo* U1 binding site in the *F9* 3'UTR could have at least two distinct outcomes: aberrant splicing or failure of 3' end processing and RNA degradation. Here, we show that our *F9* variant generates a U1 binding site. Recruitment of U1snRNP does not initiate splicing, but rather leads to suppression of the downstream PAS and to a strong decrease in *F9* mRNA

levels. This phenotype is the primary cause for reduced FIX protein levels and lower coagulation activity. Using different antisense-oligonucleotide based strategies; we rescued the mutant phenotype, which again points to U1snRNP as the key player in the pathomechanism. Furthermore, we opened the venue to a gene therapeutic approach to target such UTR variants. Our study strengthens the importance of non-coding variants as disease-causing and demands their pathomechanistic analysis.

Results

The factor 9 variant in the 3'UTR reduces *F9* mRNA expression

The previously described A to G transition at +1157 bp (c.2545A>G) downstream of the *F9* stop codon may create a 5'SS by comparison to the consensus sequence [6]. Based on our results with a similar disease-causing mutation in the *LAMTOR2* gene [12], we thought to revisit this particular variant. The 3'UTR variant is located upstream of the polyadenylation signal (PAS; Fig 1A). Bioinformatic analysis revealed a putative 3'SS downstream of the PAS in a similar distance followed by a second PAS (Fig 1A). Thus, we were left with five possible explanations: i) the variant creates a miRNA binding site; ii) splicing deletes PAS1 with insufficient PAS2 usage; iii) splicing in the 3'UTR more than 50 nts downstream of the *F9* stop codon eliciting non-sense mediated RNA decay (NMD; [13]; iv) the variant creates or deletes a motif for an RNA-binding protein (RBP) and v) the site binds U1snRNP thereby preventing PAS recognition and 3' end processing [14].

First, we generated a minigene consisting of the *F9* cDNA followed by the 3'UTR and downstream genomic sequences (Fig 1A). The mutated sequence shows high complementarity to U1 snRNA with the A to G transition creating the invariant GU of spliceable U2 introns (Fig 1A; [9]). Next, we generated the variant and two other versions either increasing (FIXopt) or decreasing (FIXdown; Fig 1A) sequence identity to the U1snRNA consensus as expressed by the number of possible base pairs to U1snRNA and the MaxEnt score describing splice site strength [15]. The variant creates a high scoring U1 binding site (Fig 1A).

A comparison of *F9* mRNA levels indicates a dramatic decrease of FIX expression in HEK 293T cells transfected with FIXmut (Fig 1B, lanes 3 and 4). Thus, the variant exerts its phenotype by downregulating *F9* mRNA amount. Suppression of the FIX PAS causes read-through into the plasmid backbone (Fig 1B and 1D). Notably, neither optimization (FIXopt; Fig 1B, lane 5) nor deoptimization (FIXdown; Fig 1B, lane 2) of the sequence did alter *F9* mRNA levels in a significant manner compared to the mutant or the wild type, respectively. To analyze this in more detail and to definitely exclude miRNAs as causative factors, we cloned another version containing the variant combined with low complementarity to U1snRNA (mutdown, S1A Fig). Although this construct carries the variant, the inability to bind U1 prevents RNA decrease (S1B Fig). This demonstrates the relevance of the mutated position for U1snRNP recognition transforming a sequence motif into a strong U1 binding site causing FIX RNA decrease.

The *F9* gene holds a potential 3'SS 198 nts downstream of its PAS (Fig 1A). This can give rise to a splicing event initiated from the newly created 5'SS and thereby removing the canonical PAS leading to nonsense-mediated RNA decay [13] or to a general failure in 3' end processing due to an insufficient PAS2. To examine these possibilities, a 3'Rapid Amplification of cDNA Ends (3'RACE) was conducted (Figs 1A and S1C). Sequencing indicates the absence of a splicing event, since the gene-specific forward primer is located upstream of the potential 5'SS (S1C Fig). In addition, the 3'RACE revealed appropriate polyadenylation at the annotated site in both FIXwt and FIXmut transfected cells (Fig 1C upper and lower panel). Importantly, these results show that the residual, non-read-through mutant *F9* mRNA (Fig 1B, lane 4) still underwent 3' end processing at the *F9* PAS1. Polyadenylation at PAS2 occurred neither in the wildtype and nor in the mutant. To excluded splicing by a more direct approach we performed

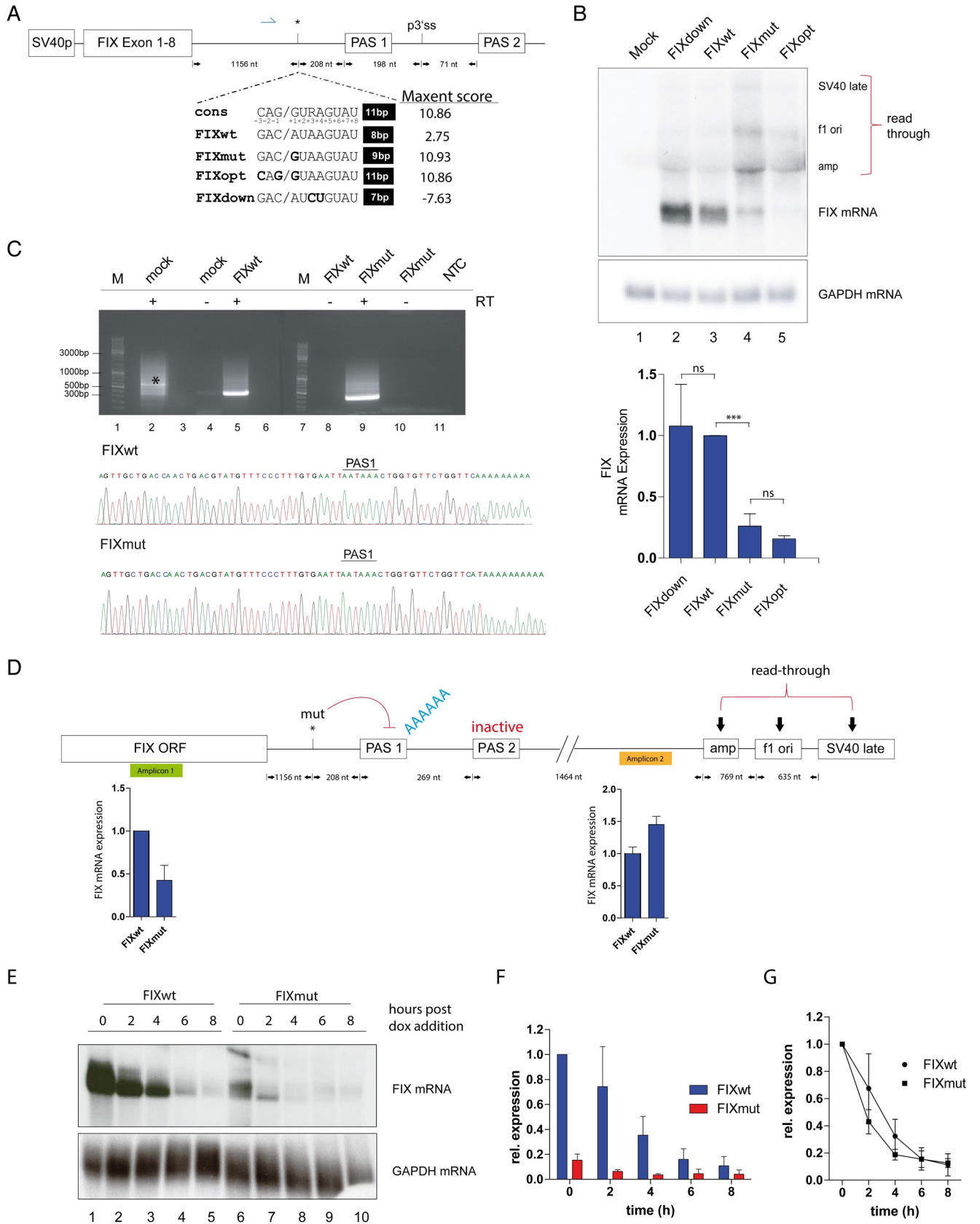


Fig 1. The variant leads to lower *F9* RNA levels and creates a putative 5' splice site. A) Depiction of the SV40 driven Factor IX minigene. The *F9* cDNA (Exon 1–8) is followed by its authentic 3'UTR. The position of the primer used for the 3'RACE is marked by a blue arrow. The variant is indicated by an asterisk at position 1156 downstream of the stop codon and 208 nts upstream of the polyadenylation site 1 (PAS1). Further downstream a putative 3'SS (p3'ss) as well as a potential second polyadenylation site (PAS2) was annotated. Another 421 nts of genomic sequence downstream of PAS2 were included. The sequence of the putative U1 binding site (5'SS) (FIXmut) is opposed to the corresponding regions of the wildtype and the 5'SS consensus sequence. The oblique outlines the exon/intron border within a common 5'SS. The numbers of possible base pairs to U1 snRNA are shown on the right, followed by the respective MaxEnt score. The position of each nucleotide within the consensus sequence is indicated by numbering. In addition, we cloned minigenes harboring either an optimized (FIXopt) or no U1 binding site (FIXdown). B) Northern blot using total RNA obtained from HEK 293T cells transfected with the indicated constructs. The membrane was hybridized with a P³²-labelled probe corresponding to the *F9* cDNA. The position of *F9* mRNA is indicated on the right. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) serves as a loading control. Bands were densitometrically quantified using a phosphoimager (lower panel). The wildtype was set to 1. The standard deviation represents 3 independent experiments. An unpaired student's T-test was performed (ns, non-significant; *** p<0.0001). To monitor transfection efficiency all experiments included transfection of a eGFP encoding plasmid. C) Agarose gel electrophoresis of 3'RACE products. Two gels are shown with markers on the left side. The presence of reverse transcriptase (RT) during cDNA synthesis is indicated (+/-). The band around 300bp represents the amplicon with the forward primer shown in (A). The signal marked with an asterisk in the untreated sample (mock) represents an amplicon of the *MLX* gene. In the lower panel Sanger sequencing of *F9* specific amplicons showed polyadenylation at the correct *F9* PAS1. D) Schematic of *F9* ORF, 3'UTR and downstream sequences including ampicillin sequence, fl ori and SV40 poly (A) signal. Suppressive effect of the variant (marked by an asterisk) on PAS1 is indicated in red. Green box symbolizes the qRT-PCR amplicon 1 detecting all *F9* RNA species and the yellow box indicates amplicon 2 for measuring *F9* RNA species derived from read-through. qRT-PCR results are depicted below the corresponding amplicons. Each qRT-PCR experiment was conducted in biological and technical triplicates. E) Tet-off *F9* RNA half-life experiment. pTetbi FIXwt and pTetbi FIXmut plasmids were transfected into HeLa TA cells. 36 hours post transfection, doxycycline was added to shut off transcription. Total RNA was collected 0, 2, 4, 6 and 8 hours post shut off and analyzed by Northern blot (n = 3). F) Quantification of *F9* RNA obtained from (E). FIXwt was set to 1 and all FIX bands were GAPDH normalized. G) Depiction of relative *F9* RNA decay rate, both FIXwt and FIXmut were set to 1 to allow direct comparison.

<https://doi.org/10.1371/journal.pgen.1008690.g001>

RT-PCR with primers located upstream of the potential 5'SS and downstream of the putative 3'SS. Again, no splicing could be detected (S1D Fig). Thus, the lack of splicing from the new, cryptic 5'SS excludes NMD and processing at PAS2 as pathomechanisms.

The PAS inhibition exerted by U1 is only transient and spatially limited to roughly 1 kb [16, 17] and allows processing at downstream sites as shown in Fig 1B. To quantify the extent of read-through we used qRT-PCR with one amplicon detecting all *F9* mRNA and a second measuring only the read-through (Fig 1D). The variant increased read-through (amplicon 2, Fig 1D, right side) as already observed in the Northern blot analysis (Fig 1B). The downregulation of mutant *F9* mRNA is not as pronounced, because all RNAs including the read-through are detected (amplicon 1, Fig 1D, left side).

The residual RNA expression observed in Fig 1B and 1D could be an artifact of the plasmid-based transient transfection. Thus, we introduced our FIX minigene into a pre-tagged, doxycycline-inducible locus to obtain a stable, single copy *F9* HeLa cell line for the wild type and the mutant (S1E Fig). Even from this genomic localization a residual RNA expression for the mutant was detectable (S1E Fig).

The observed reduction in *F9* mRNA could be due to co-transcriptional RNA decay as described for suppression of 3' end processing [18]. Alternatively, the variant may alter the *F9* mRNA stability post-transcriptionally. To analyze this, we assessed the steady-state level of *F9* mRNA and its half-life. Using a doxycycline inducible version of the FIX minigenes, we shut off transcription 36 hours post transfection. The wild type *F9* mRNA produced from this minigene decays with a half-life of approximately 3 hours (Fig 1E and 1F). The mutant *F9* mRNA is expressed at lower levels already at time point zero (compare to Fig 1B). The half-life analysis suggests only a slightly accelerated decay (Fig 1G). Thus, the mutation does not alter the half-life of *F9* mRNA.

A modified FIX minigene phenocopies poly(A) site suppression on the protein level

Our initial attempts to detect FIX protein and to measure active FIX in the supernatant of several cell lines transfected with our minigenes failed due to low expression levels (Fig 2B, lanes 2 and 3). The usage of the CMV promoter and the 5'UTR of pcDNA3.1 (Fig 2A) slightly enhanced the *F9* mRNA expression, but still recapitulated the phenotype of the variant (Fig

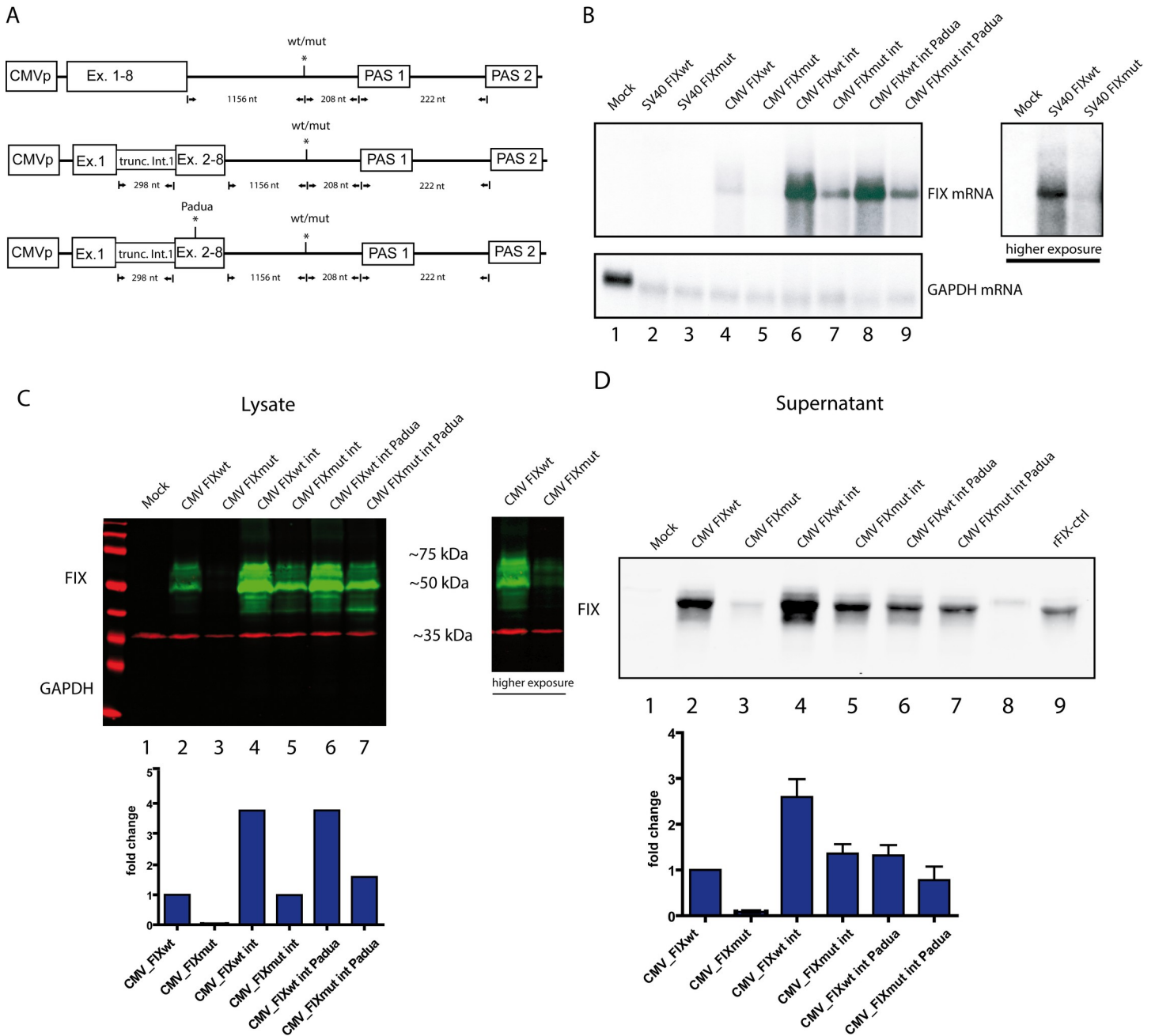


Fig 2. Establishment of a FIX protein expression system. A) Vector architecture of three different CMV driven FIX minigenes with or without 3'UTR variant. Two constructs (middle and lower) were equipped with a truncated FIX intron 1 located in between exon 1 and exon 2. The lower construct combines truncated intron 1 and the gain-of-function mutation denoted as "Padua". B) Northern blot using total RNA of HEK 293T cells transfected with the indicated constructs. GAPDH served as a loading control. A higher exposure of lanes 2 and 3 is shown in the right panel. C) Western blot of HEK 293T cell lysate transfected with the respective constructs visualized by Licor Odyssey and two antibodies labelled with different fluorophores (FIX in green). GAPDH (red) serves as loading control. Cells were lysed 30 hours post transfection. A higher exposure of lanes 1 to 3 is shown in the right panel. A densitometric quantification is depicted in the lower panel. The CMV-driven FIXwt construct was set to 1. D) Immunoblot analysis of supernatant collected from transfected HEK 293T cells with the respective constructs at 30 hours post transfection. The densitometric analysis was performed as in C. Standard deviations represent three independent experiments.

<https://doi.org/10.1371/journal.pgen.1008690.g002>

2B, lanes 4 and 5). Now we observed FIX protein both in whole cell lysates (Fig 2C, lanes 2,3 and inset) and in the supernatant (Fig 2D, lanes 2 and 3). Since FIX is synthesized in hepatocytes, we sought to verify our results in hepatic cell lines. To this end we transfected the human

hepatoma cell line Huh7 with the CMV-driven FIX constructs (S2 Fig). Also, in this cell line the variant leads to 4-fold downregulation of *F9* RNA (S2A Fig). The same is true for FIX protein in cell culture supernatants (S2B and S2C Fig).

We additionally inserted a truncated *F9* intron 1 with its authentic splice sites at the correct position into the *F9* cDNA (Fig 2A, middle panel; [19]). The intron is efficiently spliced (S3A Fig) and it strongly enhances *F9* mRNA and protein levels (Fig 2B, lane 6 and 2C, lane 4). Elevated transcription rather than splicing seems a likely explanation for this increase, since incorporation of the terminal *F9* intron did not lead to higher mRNA levels (S3B–S3D Fig) and work done by the Kurachi laboratory also implied the presence of a transcriptional enhancer [19]. Of note, the variant again led to lower RNA and protein levels even though the basal expression was higher (Fig 2B, lane 7 and 2C, lane 5). This argues for a high expressivity of the variant and a tight suppression of 3' end processing. Finally, we took advantage of a previously published gain-of-function variant referred to as Padua [20] (Fig 2A, lower panel) yielding a hyperactive FIX. Since the Padua mutation exerts its effect on FIX protein activity, the mRNA levels remained unchanged (Fig 2B, lanes 8 and 9). However, protein levels in the supernatant were slightly reduced (Fig 2D, lanes 6 and 7). Since FIX is highly modified on the post-translational level [21], we included recombinant FIX as a size-matched control showing a higher molecular weight (Fig 2D, lane 9) indicating insufficiently modified FIX protein in both the cell lysate and the supernatant (Fig 2C and 2D). Densitometric analysis revealed that the mutant phenotype gets blurred in the context of FIX intron and the Padua mutation (Fig 2D, lower panel). However, in the cell lysates a robust 4-fold decrease was observed on protein level (Fig 2C, lower panel) comparable to the reduction of *F9* mRNA (Fig 1B, lower panel).

The improved FIX minigene phenocopies the mutant also in reduced coagulation activity

CHO cells were shown to produce active FIX *in-vitro* to levels almost comparable with primary hepatocytes and they are harnessed to purify FIX for injection into hemophilic patients [22]. And indeed, transfection of the minigenes into CHO cells could reproduce the phenotype at the level of RNA (Fig 3A) and intracellular FIX protein (Fig 3B, lanes 4,5 and 6,7). Transfection of CHO cells displayed a favourable ratio of modified vs. unmodified FIX protein (compare Figs 2C and 3B). In addition, we quantified FIX antigen (FIX:Ag) by ELISA and observed a 4-fold reduction for all mutant constructs (Fig 3C). To assess coagulation activity, we employed a one-stage coagulation assay [23]. Again a 4-fold decrease in activity was observed, when comparing mutant with wildtype constructs (Fig 3D). This was also observed in hepatic Huh7 cells albeit on a lower level (S2D Fig). The Padua mutation displayed a clear gain-of-function phenotype as described previously (Fig 3D; [20]). Although protein and FIX:Ag levels were quite similar (Fig 3B and 3C), the activity was enhanced by 10 fold (Fig 3D). Again, the phenotype of our variant persisted in the hyperactive Padua context showing that the variant acts on a wide range of activity levels (Fig 3D).

In summary, the non-coding variant creates a U1 binding site in close proximity to the *F9* poly(A) site and this results in failure of 3' end processing and ultimately *F9* mRNA degradation. The decreased mRNA levels result in lower protein amounts, less secreted FIX and finally reduced FIX antigen and coagulation activity, thus partially reflecting the clinical phenotype observed in patients [24].

Blockade of U1snRNP rescues FIX expression

To demonstrate that the binding of U1snRNP to the mutated region is responsible for the decrease in *F9* mRNA, we blocked U1snRNP using Antisense Morpholino Oligonucleotides

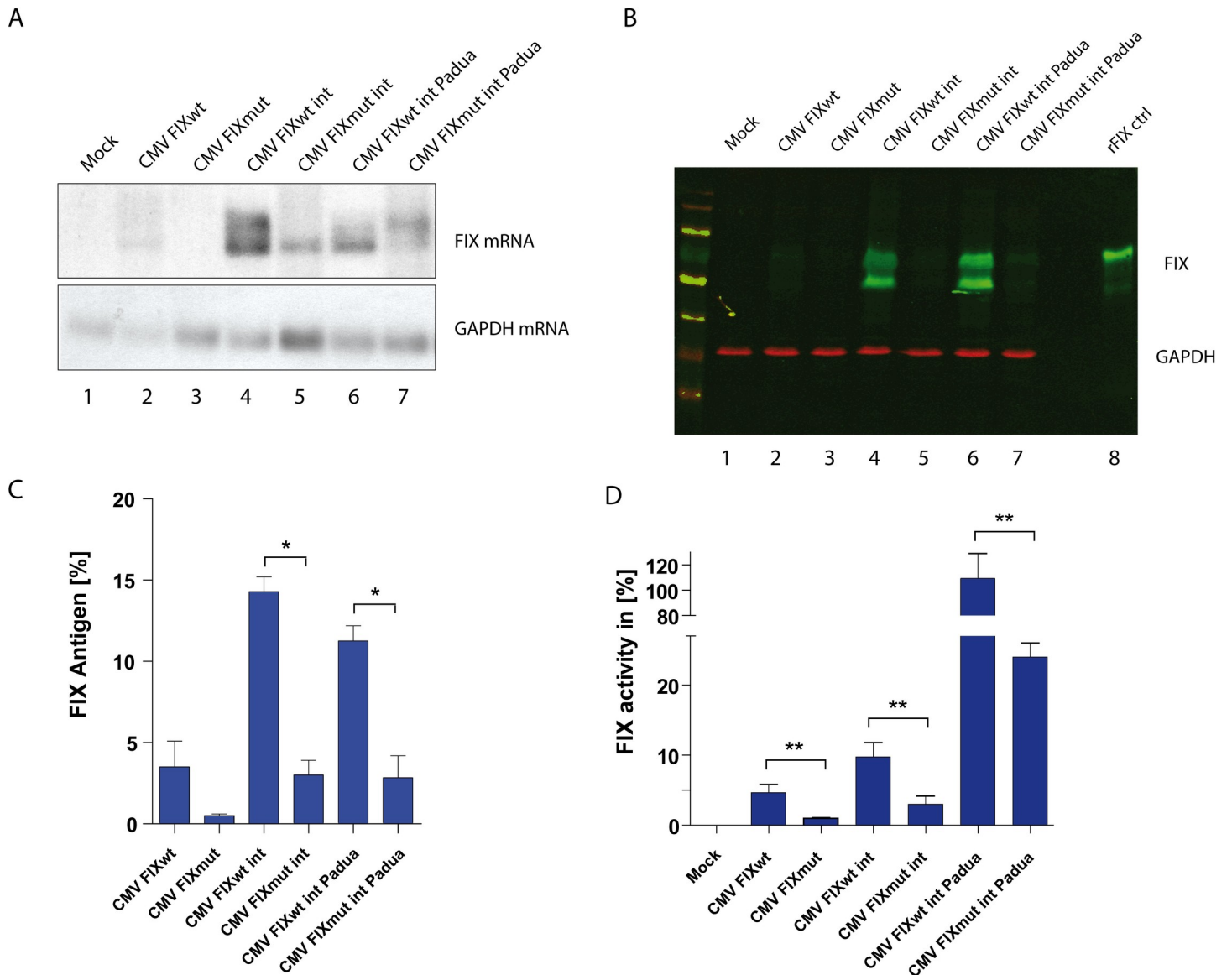


Fig 3. The F9 3'UTR variant reduces FIX antigen and coagulation activity. A) Northern blot of CHO total RNA collected 30 hours post transfection with CMV driven FIX minigenes. GAPDH serves as a loading control. B) Western blot of CHO cell lysate 30 hours post transfection with the respective minigenes. GAPDH was used as a loading control and recombinant FIX serves as a positive control (Lane 8). C) Supernatant of FIX transfected cells collected 96 hours post transfection was subjected to quantification using a FIX ELISA (n = 3). Significant differences determined via Bonferroni's Multiple Comparison Test are indicated by an asterisk ($p < 0.05$). (D) Coagulation activity of FIX transfected CHO supernatant was determined 96 hours post transfection (n = 3). Asterisks indicate statistical differences ($p < 0.05$).

<https://doi.org/10.1371/journal.pgen.1008690.g003>

(AMOs) targeting U1snRNA (Fig 4A, red sequence). U1 AMOs block 5'SS recognition by binding to the free U1 snRNA 5' end and additional nucleotides thereby invading and disrupting the secondary structure of U1snRNA [16]. We transfected the respective FIX minigenes followed by introduction of the AMOs in HeLa cells since they proved to be more resilient to the treatment. The mutant phenotype remained unaltered in the presence of a control AMO (Fig 4B, lanes 2,3 and lowest panel). The transfection of U1-specific AMOs facilitated a strong increase in mutant F9 mRNA expression by 3-fold (Fig 4B, lane 5 and lowest panel). Thus, the blockade of U1snRNP rescues the F9 mutant mRNA expression. Interestingly, both FIXwt and FIXmut minigenes showed elevated expression (Fig 4B, lanes 4,5). However, the level of

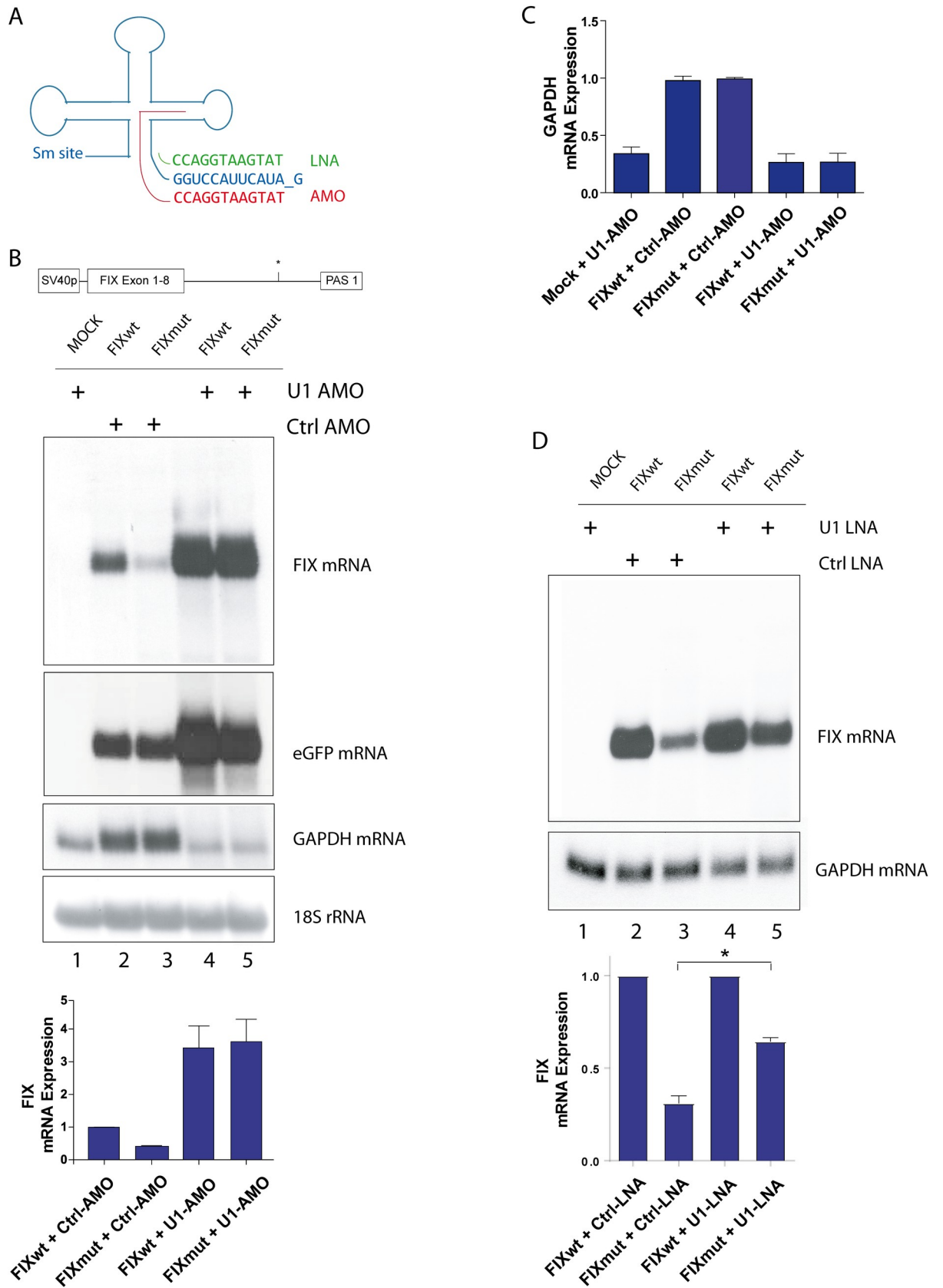


Fig 4. Blockade of U1snRNP rescues mutant F9 RNA expression. A) U1snRNA and its 5'SS recognition sequence (blue). Binding properties of locked nucleic acids (LNAs), indicated in green and antisense morpholino oligonucleotides (AMOs), indicated in red. B) Depiction of SV40-driven F9 expression construct. Below, Northern blot using total RNA from FIXwt/mut transfected HeLa cells performed as in Fig 1B. Control or anti U1 AMOs were transfected 14 h post FIX minigene transfection as indicated. The position of FIX RNA is marked on the right. The blot was re-hybridized with a GAPDH-specific, an eGFP-specific and an 18S rRNA-specific probe. Densitometric analysis of FIX mRNA expression normalized to 18S RNA is shown on the lowest panel. The FIXwt in presence of control AMOs was set to 1 (n = 3). C) Densitometric analysis of GAPDH mRNA expression. GAPDH mRNA of FIXmut transfected cells plus control AMOs was set to 1. D) Northern blot as performed in (B) except that control and anti U1 LNAs were used. The statistical analysis of three independent experiments is shown below. The asterisks indicate that the partial rescue due the blockade of U1 is significant as determined via Student's T-test.

<https://doi.org/10.1371/journal.pgen.1008690.g004>

GAPDH mRNA was decreased due to splicing inhibition upon U1snRNP blockade (Fig 4B). The 3-fold decrease in GAPDH mRNA (Fig 4C) strongly correlated with its mRNA half-life in HeLa cells of 8 hours [25]. To rule out loading differences we took advantage of 18S rRNA as an unspliced, pol I derived transcript (Fig 4B). Interestingly, expression from a GFP-encoding plasmid is also increased by anti U1 AMOs (Fig 4B). Thus, the reason for enhanced FIX expression in the wildtype seems to be a general effect of the U1 AMOs on intron-less genes. Very recently U1 AMOs were shown to increase RNA pol II density at transcriptional start sites, which may explain our observation [26].

In a second approach, a blockade of U1snRNP has been performed using locked nucleic acids (LNAs) designed exclusively against the U1snRNA 5' end (Fig 4A, green sequence), thereby retaining the U1snRNA secondary structure. Masking of the U1snRNA 5' end led to a rescue of F9 mRNA expression in FIXmut transfected cells (Fig 4D, lanes 4 and 5), whereas cotransfection with a control LNA still showed reduced FIXmut mRNA expression. The rescue was incomplete due to the lower transfection efficiency of LNAs vs. AMOs. Again, the GAPDH expression was reduced due to the manipulation of splicing. Importantly, after LNA treatment the wildtype F9 mRNA expression remained constant (Fig 4D, lanes 2 and 4).

A U1snRNA suppressor mutant reduces wild type F9 RNA expression

In a third approach we sought to co-express a suppressor U1snRNA, which is designed to bind the wildtype sequence and to a lesser extent to the variant (Fig 5A; [12]). To this end we co-transfected our FIX minigenes along with the U1 suppressor mutant plasmid or an empty plasmid as a control. The retargeting of the suppressor U1snRNP to the wildtype minigene decreases F9 mRNA expression (Fig 5B). RNA expression from the mutant minigene remains constant (Fig 5B). A rescue of the mutant RNA by the suppressor U1 snRNA is not expected since a G•U base pair still allows binding of the suppressor to mutated sequence (Fig 5A and 5B). In addition, high levels of endogenous U1snRNP are still repressing 3' end processing of mutant FIX minigene.

This approach together with the data presented above clearly demonstrates a direct involvement of U1snRNP in the suppression of 3' end processing in the context of the 3'UTR variant. In addition, we only found one miRNA binding site, which seed region overlaps with the variant but leading to a lower score (S4 Fig). Moreover, we employed the algorithm RBP map [27] to identify whether the variant changed binding motifs for RBPs. The variant, however, creates an ANKHD1 binding site, which is involved in signaling [28] and transcriptional regulation via histone binding [29] and is thus unlikely to be responsible for the observed phenotype.

Masking the mutated site rescues FIX expression

While the mechanistic approach clearly showed that the pathomechanism of the variant depends on U1snRNP recognition, blockade of U1 is not a therapeutic option. We followed a vice versa approach and aimed to mask the newly created U1 binding site by chemically

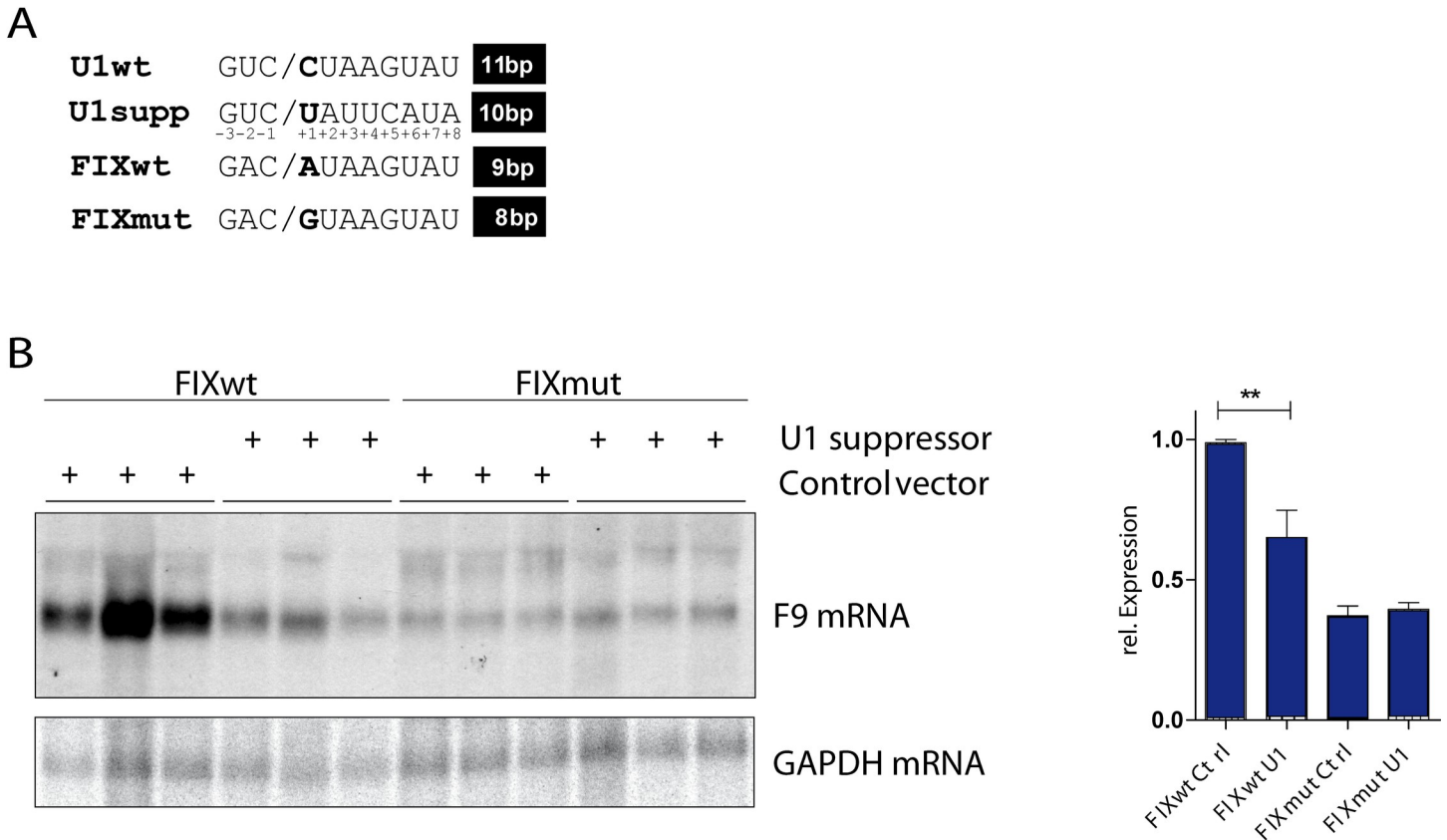


Fig 5. A U1 snRNA suppressor mutant reduces wild type F9 RNA expression. (A) The U1 suppressor mutation is depicted. In addition, the wildtype sequence (FIXwt) is opposed to the corresponding regions of the variant (FIXmut) and the 5'SS consensus sequence. The oblique outlines the exon/intron border. The number of possible base pairs to U1snRNA are shown on the right. The position of each nucleotide within the consensus sequence is indicated by numbering. (B) HEK 293T were transfected in triplicates with the indicated minigenes and U1 suppressor plasmids or control vectors. The Northern blot was hybridized with a F9 specific probe and re-hybridized with a GAPDH probe for normalization. A statistical analysis of three independent experiments is shown on the right side. ** = $p < 0.004$.

<https://doi.org/10.1371/journal.pgen.1008690.g005>

modified RNA antisense oligonucleotides (2OMePS) displaying increased stability [30] and therapeutic potential in Duchenne muscular dystrophy (DMD);[31, 32]. We designed a 20mer 2OMePS antisense oligonucleotide hybridizing to both the wildtype and the mutated sequence (Fig 6A). To this end we transfected our original SV40-driven FIX minigenes along with different concentrations of control and targeting oligonucleotides into HeLa cells, which again are more resilient to the 2OMePS treatment (Fig 6B). This vector change was necessary, since our optimized constructs produced high amounts of F9 mRNA, which could not be rescued due to a lower efficiency of 2OMePS transfection and thus, a lower nuclear antisense oligonucleotide concentration in HeLa cells. While in transfections with a control oligo the decrease of mutated F9 mRNA persisted (Fig 6B, lane 2 and 3), masking the site to prevent association of U1snRNP leads to a rescue of mutant F9 mRNA expression (Fig 6B, lane 4 and 5). Higher concentrations of the oligo exhibited some toxicity as shown by the lower GAPDH signal (Fig 6B, lane 6). Interestingly, an additional F9 mRNA species is detectable in both wildtype and mutant context indicating a switch to a processing site beyond PAS2 (Fig 1A).

The same strategy can also be applied using a viral vector derived system. Here, the snRNP system is 'hijacked' by replacing the 5' part of natural snRNAs with an antisense sequence of choice. This strategy protects the antisense sequence of choice from degradation. U7snRNA is often used for this purpose, which normally ensures histone RNA processing (Fig 6C;[33, 34].

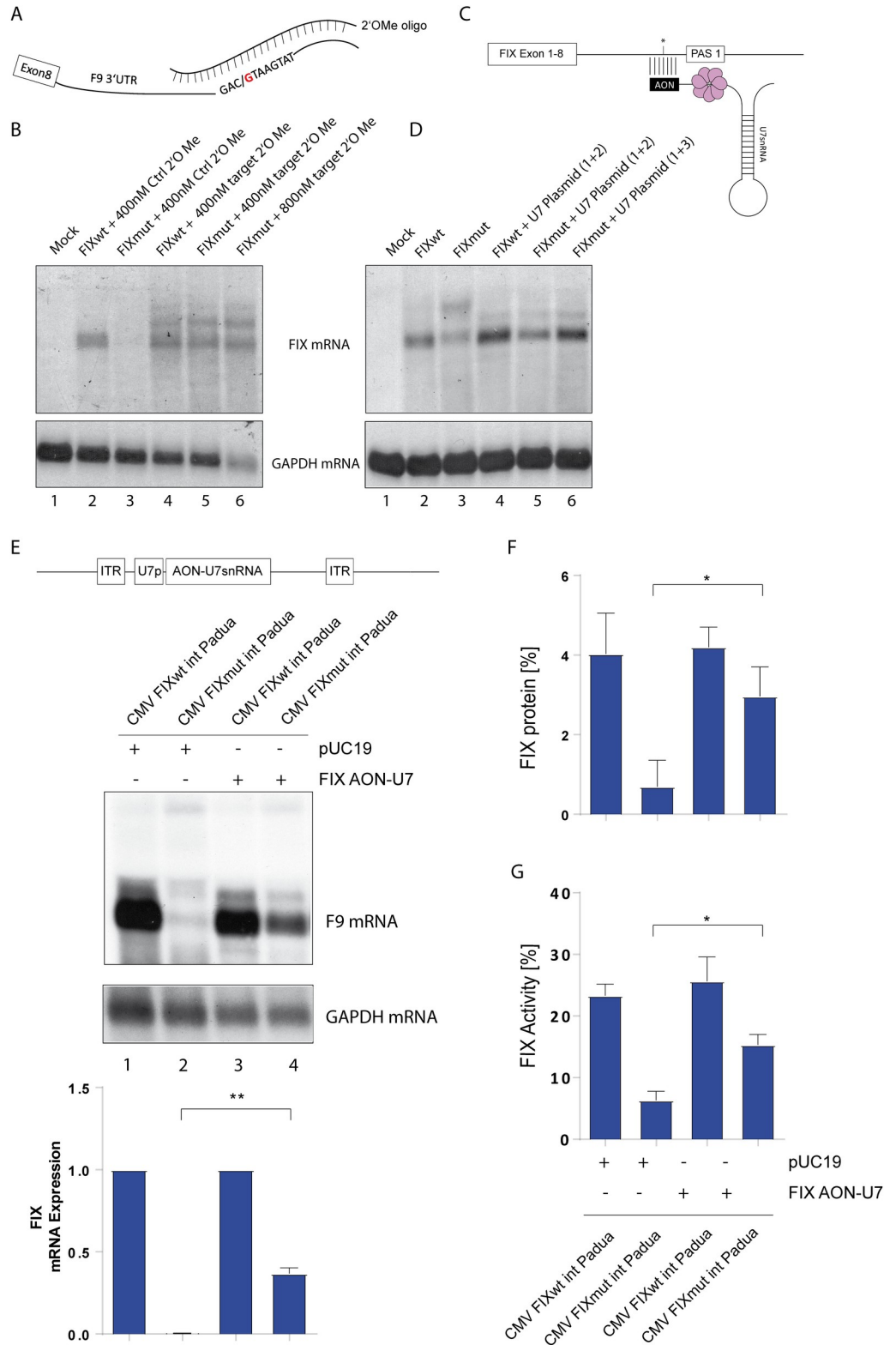


Fig 6. Masking of mutated region via two different antisense oligonucleotides. A) Depiction of the mutationally generated U1 binding site being masked by a complementary 2'-O-Methyl oligonucleotide (2OMePS). B) Northern blot of total RNA after co-transfection of SV40 driven FIXwt and mut constructs with control or targeting 2OMePS oligonucleotides. HeLa cells were harvested 16h post transfection. GAPDH serves as a loading control. C) Depiction of

the modified U7snRNP targeting the mutated region on the FIX mRNA thereby masking the U1snRNP binding site. D) Northern blot of total RNA after cotransfection of SV40 driven FIX minigenes and U7-antisense expression vector in indicated molar ratio (1+2/1+3). HeLa cells were harvested after 16 hours as in (B) and GAPDH serves as a loading control. E) Depiction of the AAV vector backbone, which was used in transient transfections (top). The inverted terminal repeats (ITR) delineate the vector-derived sequences. A cassette consisting of the U7 promoter (U7p) and a fusion of the FIX antisense sequence (AON) to U7 snRNA was inserted. The lower panel the RNA analysis by northern blot using total RNA from transfected CHO cells in presence of the AAV plasmid expressing the antisense oligonucleotide fused to U7snRNA or pUC19 as a control. GAPDH serves as a loading control. Below three independent experiments were quantified by phosphoimager analysis. The partial rescue in the presence of the antisense oligonucleotide is indicated by a double asterisk ($p = 0.0025$). F) Supernatant of FIX transfected CHO cells collected 96 hours post transfection was subjected to quantification using a FIX ELISA ($n = 3$). Significant differences determined via Bonferroni's Multiple Comparison Test are indicated by an asterisk ($p = 0.05$). G) Coagulation activity in percent matched to a standard curve of recombinant factor IX. The tested supernatants were derived from CHO cells transfected with the indicated plasmids. Standard deviations represent three independent experiments. The asterisks indicate a statistically significant elevation of FIX activity ($p < 0.05$).

<https://doi.org/10.1371/journal.pgen.1008690.g006>

We constructed a plasmid based on an adeno-associated virus (AAV) vector backbone [35, 36] and inserted a longer antisense sequence of 40 nucleotides complementary to the *F9* 3'UTR fused to U7snRNA (Fig 6C and 6E; upper scheme). This sequence shows no hits in a Blast search (S4B Fig). Transfection along with the FIX minigenes showed a rescue of mutant *F9* mRNA expression (Fig 6D, lanes 4–6). Again, a more downstream PAS is used when the U1 binding site is masked (Fig 6D, lane 6). Next, we introduced the FIX minigene and the U7 plasmid into CHO cells by transfection and monitored *F9* mRNA by Northern blot (Fig 6E, middle panel). The co-transfection of the U7 plasmid led to a significant increase in mutant *F9* mRNA expression (Fig 6E, lowest panel). This rescue is also detectable on the protein level in the supernatant measure by FIX ELISA (Fig 6F) and finally also the coagulation activity was increased by blocking U1 access to the mutant site (Fig 6G). The absolute activity was lower compared to Fig 3 due to altered molar ratios of the transfected FIX plasmid (Fig 6G). This confirms the potential of our strategy to be applied as a gene therapeutic approach to treat non-coding and disease-causing variants leading to hemophilia B in this case and other diseases.

Discussion

A reassessment of UTR variants in *F9* revealed a point mutation possibly creating a 5'SS [6]. We promisingly recapitulated the phenotype demonstrating a downregulation of *F9* mRNA in the presence of the variant. A key player of this pathomechanism is U1snRNP, which erroneously binds within the 3'UTR and leads to a PAS suppression resulting in poor mRNA expression. Thus, the variant misleads the U1-dependent surveillance mechanism, which normally suppresses intronic poly(A) sites [37] and a failure of 3' end processing results in RNA degradation [18].

The *F9* mRNA repression also translates into reduced FIX protein levels explaining the hemophilia B observed in the patients (<http://www.factorix.org>; [8]). While we clarified the pathomechanism, our minigene system is not able to perfectly mirror the reduction of FIX antigen and coagulation activity observed in the patients. This is mainly due to a higher residual mutant RNA expression. In addition, the read-through RNAs (Fig 1B and 1D) may also contribute to FIX protein, thus slightly compensating the effect of the variant. Also, a human hepatic cell line displayed the same residual RNA expression (S2 Fig) and so did a single integrated copy of *F9* under control of a dox-inducible promoter (S1C Fig). We speculate that the genomic location on the X chromosome in conjunction with local regulatory sequences is the reason for this discrepancy. In addition, we always assayed for accumulation of FIX in the supernatant compared to secretion and circulation/consumption of FIX in the patients.

Nevertheless, our *F9* minigene proved to be a useful tool to uncover the pathomechanism of the non-coding variant.

Despite the different levels of transcription, the phenotype of the variant persisted, which again stresses its high expressivity also shown by the reduced coagulation activity of our mutant supernatants. At the molecular level this can be explained by the high excess of U1snRNP in the nucleus leading to a tight suppression of 3' end processing. Due to issues on FIX modifications and activity during *in-vitro* experiments, we sought to use the Padua mutation as a “magnifier” to thoroughly test the expressivity of our non-coding variant. Now the coagulation assay reached the same activity as recombinant FIX and the 4-fold downregulation of the variant persisted also in this context.

We reasoned that shielding of the mutant site would prevent U1snRNP-binding and alleviate PAS suppression. Two types of antisense oligonucleotide (ASO) strategies previously demonstrated efficacy in other models: i) 2OMePS RNA oligos [32, 38] with some side effects reported [39] and ii) an plasmid based system expressing an antisense oligo fused to U7 snRNA [35]. In essence, both antisense strategies lead to a rescue of *F9* mRNA expression and an increase in coagulation activity, which may be able to revert the phenotype into a mild or subclinical hemophilia B. Interestingly, the RNA duplex formation and association with U7snRNP did not hamper translation in the cytosol, which suggests detachment of the complex at the nuclear pore or during the pioneer round of translation. However, we will have to determine if the partial rescue was due to limitation of the cellular processing capacity of snRNP particles [40].

Most hemophilia B patients suffer from disease-causing variants in the coding region (<http://www.factorix.org>). Some of these patients have been treated by gene-addition therapy introducing a non-mutated *F9* cDNA as it has been done using adeno-associated viral vectors (AAV) in clinical trials [41]. We constructed similar vectors expressing the antisense-U7snRNA cassette as mentioned above and produced recombinant AAV particles. However, cell culture experiments using AAV serotype 2 did not reach sufficient antisense RNA levels to reproduce the partial rescue of FIX activity observed by transfection of the AAV vector plasmid (Fig 5E). We speculate that the U7 hairpin structure may serve as a resolving signal for the AAV rep protein leading to the aberrant packaging of truncated genomes. Currently we are optimizing the vector design. In general, this current study might be a starting point for dissecting and treating other UTR variants, which exert their effects by misplacing cellular RNA-binding proteins on the mutant RNAs. These variants may cover a wide range of phenotypes from modulation of RNA half-life to mislocalization in the cytosol. For example, the Factor VIII mutation database list four 3'UTR variants. One advantage of the variant masking strategy compared to gene addition is the ongoing physiological expression of the mutated gene as only transcribed RNAs are targeted and rescued.

Materials and methods

Expression vectors

The human *F9* cDNA (NM_000133.3) was cloned into pGL3 (Promega). The 3'UTR including the variant and 597 bp of downstream sequence was synthesized by MWG operon.

FIXwt was restored by PCR mutagenesis (see primer list in Table 1). FIXdown and FIXopt were generated by PCR with primers P9, P17 and P18 and cloned into pGL3 using EcoRI and XbaI. The 5'UTR of pcDNA3.1 and a CMV promoter were inserted using BglII and HindIII. A sequence comprising Factor IX Padua and a truncated *F9* intron 1 was synthesized by MWG operon and inserted into the CMV driven factor IX minigene via HindIII and BamHI. The Padua mutation has been replaced by the authentic sequence using BstBI and MfeI in a partial

digest. The vector coding for U7snRNA linked to a mutated factor IX complementary region descends from pSMD2 U7dtx23 [42]. F9 specific antisense nucleotides were inserted into the U7snRNA gene by three PCR reactions (P8+P10, P9+P11 and P10+P11) and implemented into pSMD2 GFP via XbaI/NheI. F9 vectors employed for the decay experiments and stable cell lines are derived from CMV FIXwt int Padua or CMV FIXmut int Padua. F9 cDNA sequence and 3'UTR were amplified using P16+P17 and transferred into a pTetbi promoter driven vector (kindly provided by V. Cordes, MPI Göttingen) via AgeI, BamHI and NcoI digest. FIX mutdown was generated via PCR mutagenesis using primers P1, P18, P19 and P20. Products of PCR1 (P1+P18) and PCR2 (P19 and P20) were employed for a fusion PCR using P1 and P20. The corresponding PCR product was cloned into FIXwt via EcoRI/XbaI, thereby replacing the wild type sequence and giving rise to FIX mutdown. Cloning of the U1 suppressor mutant was done by modifying the 5'SS recognition site of the U1snRNA to bind „GACA-TAAGT“. This modified U1snRNA including the authentic U1snRNA promoter was synthesized (IDT-DNA) and cloned into the U1 wt vector [12] via HindIII/BamHI.

Cell culture and transfection

HEK 293T, HeLa and CHO cells were grown in DMEM supplemented with 10% fetal calf serum, 1mM sodium pyruvate, and 1% penicillin/streptomycin and 4μg/mL vitamin K at 37°C and 5% CO₂. One day before transfection, 5×10⁶ 293T cells were seeded in a 10-cm dish to a confluency of 70–80%.

Transfections were performed using calcium phosphate and a total DNA amount of 15.5μg (5μg FIX plasmid, 0.5μg eGFP DNA and 10μg of pUC19 DNA). Cells were harvested 16h post transfection for RNA analysis. pTetbi FIXwt or pTetbi FIXmut were co-transfected with a Flp-recombinase vector into Hela-EM2-11ht cells [43]. Single cell clones were evaluated via Mammalian beta-Galactosidase Assay Kit (ThermoFisher) according to the manufacturer's protocol. Half-life experiments were conducted in HeLa TA cells harboring a stably integrated copy of the Tet-transactivator by the addition of 20 ng Doxycyclin. For U1 blockade, HeLa cells were transfected with FIXwt and FIXmut using Viafect (Promega). 18 hours post transfection, cells were PBS-washed and trypsinised. After centrifugation, 10⁶ cells were resuspended in 50μL of Buffer "R" (Neon Transfection System, Life Technologies) and 2μL of 300nmol of antisense morpholino oligonucleotides (AMOs) were added. The cell suspension was successively exposed to two pulses of 35ms width and an input voltage of 1005 V. After electroporation, cells were transferred in 2mL of antibiotic-free medium and incubated overnight before harvesting. For U1 blockade via locked nucleic acids (LNAs) 5×10⁵ HeLa cells were seeded in a six-well. A total DNA amount of 2μg was transfected using Viafect (Promega). 4 hours post transfection, 10nmol of anti U1 or control LNAs (Exiqon) were transfected using Lipofectamine 3000. Cells were harvested 20 hours post LNA transfection. For the U1 suppressor experiments, HEK 293T cells were seeded into 6-well plates. At a confluency of approximately 70%, cells were transfected with 0.7ug FIXwt or FIXmut and 1.8ug of either U1 suppressor vector or empty plasmid (pUC19) using Lipofectamine 3000. RNA was harvested 30 hours post transfection and subjected to Northern blot analysis.

For FIX activity analysis, 4×10⁵ CHO cells were seeded in a six-well. At a confluency of about 50%, cells were transfected with various FIX minigenes using Viafect (Promega). 12 hours post transfection, medium was replaced with FCS-free DMEM, supplemented with 4μg/mL vitamin K, 2.5% BSA and 1% L-Glutamine and cells were incubated for 96 hours.

For the 2OMePS mediated rescue, 5×10⁵ HeLa cells were seeded into a six-well one day before transfection. A total DNA amount of 2μg was transfected using Viafect (Promega). 4 hours post viafection, 400nmol of targeting (2OMePS) or fluorescent control oligo were

Table 1. List of oligonucleotides.

Cloning of <i>F9</i> cDNA	P1 (EcoRI for)	5'-GAGGAAGAATTCAACAGTGTGTCTTCAGC-3'
	P2 (FIX_opt_rev)	5'-GGAGACATGATACTTACCTGCTCTGGTC-3'
	P3 (FIX_down_rev)	5'-GGAGACATGATACAGATGTCCTCTGGTC-3'
Cloning of <i>F9</i> downstream	P4 (FIX_ovl_DS_for_wt)	5'-GTTGAAGTTGCCTAGACCAGAGGACATAAG-3'
	P5 (FIX_ovl_US_rev_wt)	5'-GTATGCTAGTTAAAGGAGACATGATACTTATGTCCTC-3'
Sequences	P6 (3'UTR BamHI for)	5'-GCGGATCCTGAAAGATGGATTTCCAAGGTTAATTCATTG-3'
	P7 (FIX_ovl_flank_DS_rev)	5'-GAGCATTGAGAAAGCGCCACGC-3'
Cloning of U7 construct	P8 (U7-FIX-Reverse)	5'-AGAGGACGTAAGTATCATGTCTTTGCGGAAGTGCCTGTGA-3'
	P9 (U7-FIX-Forward)	5'-TTACGTCCTCTGGTCTAGGAATTTTTGGAGCAGGTTTTCT -3'
	P10 (FU7-Xba)	5'-GGGTCTAGATAACAACATAGGAGCTGTGA-3'
	P11 (RU7-Nhe)	5'-AAAGCTAGCCACAACGCGTTTCCTAGGA-3'
Amplicon 1	P12 (<i>F9</i> _ORF_qPCR_for)	5'-ATTCTATGAATGTTGGTGTCCCT-3'
	P13 (<i>F9</i> _ORF_qPCR_rev)	5'-GGGTGCTTTGAGTGATG TTATCCAA-3'
Amplicon 2	P14 (<i>F9</i> _readthrough_for)	5'-GTGCACCTATAATCCAGCTACTGGGGAG-3'
	P15 (<i>F9</i> _readthrough_rev)	5'- CAGTCATAAGTGCGGCGACG-3'
Stable cell line	P16 (cl- <i>F9</i> NcoI_for)	5'- GCGCGCCATGGCAAGCTTACCACCTTCACAATCTGCTAGC-3'
	P17 (cl- <i>F9</i> AgeI_rev)	5'-CGCGACCGGTGCGAATTGGGATGCCTCTCCATG-3'
FIXmutdown	P18 (FIX mut down rev)	5'-GGAGACATGATTAAGACGACCTCTGGTC-3'
	P19 (FIX mut down for)	5'-GTCGTCTTAATCATGTCTCCTTTAAC-3'
	P20 (Xba rev)	5'- ACTCTCTAGAAGATTCAAGATAGAAG-3'
	P21 (cr ds primer)	5'- CTGGGCCAGCCAAGAAATTTAA-3'
Antisense	AMO oligonucleotide—Control	5'-CTTCTTACCTCAGTTACAATTTATA-3'
Morpholinos	AMO oligonucleotide—anti U1	5'-GGTATCTCCCCTGCCAGGTAAGTAT-3'
LNAs	Anti U1 Locked Nucleic Acid	5'-G+CC+AG+GT+AA+GT+AT-3'
	Control Locked Nucleic Acid	5'-G+CC+AA+CT+CA+CT+AT-3'
2OMePs	2'O Methyl anti FIX oligo	5'-UGAUACUUACGUCCUCUGGU-3'
Oligos	2'-O-Methyl control oligo	5'-AAAAGAAAACAUAUCACAAAAUGGG-3'

<https://doi.org/10.1371/journal.pgen.1008690.t001>

transfected using 5 μ L of Lipofectamine 3000 (Invitrogen). RNA was harvested 16 hours after second transfection. U7 mediated rescue was performed via cotransfection of SV40 driven FIXwt or FIX mut and pSMD2-FIX-U7 into HeLa cells in a molar ratio of 1:3 using viafect. Medium was changed after 4 hours and 16h post transfection; cells were harvested for RNA isolation.

RNA preparation and analysis

RNA methods were performed as described previously [12]. For detection of *F9* RNA, a specific probe corresponding to the *F9* cDNA was generated from the FIX plasmid by HindIII/BamHI digestion. The GAPDH-specific probe was prepared as described [12]. cDNA synthesis for quantitative PCR (qPCR) was conducted using QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's protocol. Prior to reverse transcription, RNA was treated with TURBO DNase (Invitrogen) and purified via RNeasy columns (Qiagen). qPCR was performed with QuantiTect SYBR Green PCR Kit (Qiagen) using primer pairs P12+P13 (amplicon 1) and P14+P15 (amplicon 2). For 3'RACE experiments, poly(A)⁺ -RNA of transfected cells was isolated and 50 ng were reverse transcribed. To this, the Invitrogen 3'-RACE protocol and the provided reagents were utilized. The PCR products were separated via agarose gel electrophoresis. After gel extraction, the DNA was transferred into pCR2.1 (Invitrogen) for sequencing.

Western blot analysis and factor IX activity measurement

For detection of FIX protein, the F9 monoclonal antibody (M01), clone 2C9 (Abnova) was employed in a dilution of 1:500. GAPDH protein expression was detected using GAPDH antibody (1:2000; Cell Signaling Technology). Detection was performed via Licor Odyssey infrared imaging system. Recombinant FIX (BeneFix, Pfizer) was employed as positive control. Supernatant obtained from CHO cells was employed for ELISA-based quantification of FIX antigen (FIX:Ag) using ZYMUTEST Factor IX from HYPHEN BioMed according to the manufacturer's protocol including a standard FIX calibrator concentration set to 100% which equals 100 U/dl. Thus, FIX:Ag levels are expressed as % FIX protein. FIX activity was determined using a one-stage coagulation assay. Samples were diluted 1:5 in imidazol buffer and mixed with an equal amount of FIX deficient plasma and Actin FS APTT reagent (all from Siemens Healthcare). After incubation at 37°C for 2 min, 0.025 M calcium-chloride was added, and the coagulation time recorded in an Amelung KC10 coagulometer. Calibration curves were generated with human standard plasma diluted in FIX deficient plasma.

Bioinformatic analysis and statistics

MicroRNA binding sites were analyzed using the online prediction tool miRDB [44]. RBPmap [27] is available here: <http://rbpmap.technion.ac.il/>. Splice site prediction was performed using NNSPLICE 0.9 version [45] and MaxEnt score [15]. Statistical analysis (Bonferroni's Multiple Comparison test) was performed using the GraphPad Prism software.

Supporting information

S1 Fig. Mutational analysis and stable FIX cell line.

(PDF)

S2 Fig. Analysis of the phenotype in the hepatic cell line HuH-7.

(PDF)

S3 Fig. Splicing of the first truncated and the terminal FIX intron.

(PDF)

S4 Fig. miRNA analysis and BLAT screening for off-target effects.

(PDF)

S1 Table. Clinical and phenotypic data of hemophilia B patients carrying the same mutation.

(PDF)

Acknowledgments

We thank Thomas Schulz for ongoing encouragement.

Author Contributions

Conceptualization: Andreas Tiede, Jörg Langemeier, Jens Bohne.

Formal analysis: Sonja Werwitzke, Dirk Varnholt, Amelie S. Wachs.

Funding acquisition: Jens Bohne.

Investigation: Simon Krooss, Jörg Langemeier, Jens Bohne.

Methodology: Johannes Kopp, Alice Rovai, Dirk Varnholt, Andreas Tiede, Jörg Langemeier.

Project administration: Jens Bohne.

Resources: Sonja Werwitzke, Amelie S. Wachs, Aurelie Goyenville, Annemieke Aarstma-Rus, Michael Ott, Andreas Tiede, Jens Bohne.

Supervision: Jörg Langemeier, Jens Bohne.

Validation: Jens Bohne.

Visualization: Jens Bohne.

Writing – original draft: Simon Krooss, Jörg Langemeier, Jens Bohne.

Writing – review & editing: Sonja Werwitzke, Dirk Varnholt, Amelie S. Wachs, Aurelie Goyenville, Annemieke Aarstma-Rus, Michael Ott, Andreas Tiede, Jörg Langemeier, Jens Bohne.

References

1. Manning KS, Cooper TA. The roles of RNA processing in translating genotype to phenotype. *Nat Rev Mol Cell Biol.* 2017; 18(2):102–14. <https://doi.org/10.1038/nrm.2016.139> PMID: 27847391; PubMed Central PMCID: PMC5544131.
2. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell.* 2017; 169(7):1177–86. <https://doi.org/10.1016/j.cell.2017.05.038> PMID: 28622505; PubMed Central PMCID: PMC5536862.
3. Anson DS, Choo KH, Rees DJ, Giannelli F, Gould K, Huddleston JA, et al. The gene structure of human anti-haemophilic factor IX. *EMBO J.* 1984; 3(5):1053–60. PMID: 6329734; PubMed Central PMCID: PMC557470.
4. Mannucci PM, Tuddenham EG. The hemophilias—from royal genes to gene therapy. *N Engl J Med.* 2001; 344(23):1773–9. <https://doi.org/10.1056/NEJM200106073442307> PMID: 11396445.
5. Rallapalli PM, Kembal-Cook G, Tuddenham EG, Gomez K, Perkins SJ. An interactive mutation database for human coagulation factor IX provides novel insights into the phenotypes and genetics of hemophilia B. *J Thromb Haemost.* 2013; 11(7):1329–40. <https://doi.org/10.1111/jth.12276> PMID: 23617593.
6. Vielhaber E, Jacobson DP, Ketterling RP, Liu JZ, Sommer SS. A mutation in the 3' untranslated region of the factor IX gene in four families with hemophilia B. *Human molecular genetics.* 1993; 2(8):1309–10. <https://doi.org/10.1093/hmg/2.8.1309> PMID: 8401514.
7. Chen SH, Schoof JM, Weinmann AF, Thompson AR. Heteroduplex screening for molecular defects in factor IX genes from haemophilia B families. *British journal of haematology.* 1995; 89(2):409–12. <https://doi.org/10.1111/j.1365-2141.1995.tb03319.x> PMID: 7873393.
8. Awidi A, Alhattab D, Bsoul N, Magablah A, Mefleh R, Dweiri M, et al. FIX mutation spectrum in haemophilia B patients from Jordan: identification of three novel mutations. *Haemophilia: the official journal of the World Federation of Hemophilia.* 2011; 17(1):162–3. <https://doi.org/10.1111/j.1365-2516.2010.02365.x> PMID: 20695909.
9. Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell.* 2009; 136(4):701–18. <https://doi.org/10.1016/j.cell.2009.02.009> PMID: 19239890.
10. Roca X, Sachidanandam R, Krainer AR. Determinants of the inherent strength of human 5' splice sites. *Rna.* 2005; 11(5):683–98. <https://doi.org/10.1261/ma.2040605> PMID: 15840817.
11. Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, et al. U1 snRNP determines mRNA length and regulates isoform expression. *Cell.* 2012; 150(1):53–64. Epub 2012/07/10. S0092-8674(12)00650-2 [pii] <https://doi.org/10.1016/j.cell.2012.05.029> PMID: 22770214; PubMed Central PMCID: PMC3412174.
12. Langemeier J, Schrom EM, Rabner A, Radtke M, Zychlinski D, Saborowski A, et al. A complex immunodeficiency is based on U1 snRNP-mediated poly(A) site suppression. *EMBO J.* 2012; 31(20):4035–44. Epub 2012/09/13. <https://doi.org/10.1038/emboj.2012.252> emboj2012252 [pii]. PMID: 22968171; PubMed Central PMCID: PMC3474926.
13. Kurosaki T, Maquat LE. Nonsense-mediated mRNA decay in humans at a glance. *J Cell Sci.* 2016; 129(3):461–7. <https://doi.org/10.1242/jcs.181008> PMID: 26787741; PubMed Central PMCID: PMC4760306.

14. Langemeier J, Radtke M, Bohne J. U1 snRNP-mediated poly(A) site suppression: Beneficial and deleterious for mRNA fate. *RNA Biol.* 2013; 10(2). Epub 2013/01/18. 23314 [pii]. <https://doi.org/10.4161/rna.23314> PMID: 23324605.
15. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of computational biology: a journal of computational molecular cell biology.* 2004; 11(2–3):377–94. <https://doi.org/10.1089/1066527041410418> PMID: 15285897.
16. Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, et al. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature.* 2010; 468:664–8. <https://doi.org/10.1038/nature09479> PMID: 20881964.
17. Shi J, Deng Y, Huang S, Huang C, Wang J, Xiang AP, et al. Suboptimal RNA-RNA interaction limits U1 snRNP inhibition of canonical mRNA 3' processing. *RNA Biol.* 2019; 16(10):1448–60. <https://doi.org/10.1080/15476286.2019.1636596> PMID: 31242075; PubMed Central PMCID: PMC6779394.
18. Chiu AC, Suzuki HI, Wu X, Mahat DB, Kriz AJ, Sharp PA. Transcriptional Pause Sites Delineate Stable Nucleosome-Associated Premature Polyadenylation Suppressed by U1 snRNP. *Mol Cell.* 2018; 69(4):648–63 e7. <https://doi.org/10.1016/j.molcel.2018.01.006> PMID: 29398447.
19. Kurachi S, Hitomi Y, Furukawa M, Kurachi K. Role of intron I in expression of the human factor IX gene. *J Biol Chem.* 1995; 270(10):5276–81. <https://doi.org/10.1074/jbc.270.10.5276> PMID: 7890639.
20. Simioni P, Tormene D, Tognin G, Gavasso S, Bulato C, Iacobelli NP, et al. X-linked thrombophilia with a mutant factor IX (factor IX Padua). *N Engl J Med.* 2009; 361(17):1671–5. <https://doi.org/10.1056/NEJMoa0904377> PMID: 19846852.
21. Hansson K, Stenflo J. Post-translational modifications in proteins involved in blood coagulation. *J Thromb Haemost.* 2005; 3(12):2633–48. <https://doi.org/10.1111/j.1538-7836.2005.01478.x> PMID: 16129023.
22. Enjolras N, Perot E, Le Quellec S, Indalecio A, Girard J, Negrier C, et al. In vivo efficacy of human recombinant factor IX produced by the human hepatoma cell line HuH-7. *Haemophilia: the official journal of the World Federation of Hemophilia.* 2015; 21(4):e317–21. <https://doi.org/10.1111/hae.12688> PMID: 25981983.
23. Rodriguez MH, Enjolras N, Plantier JL, Rea M, Leboeuf M, Uzan G, et al. Expression of coagulation factor IX in a haematopoietic cell line. *Thrombosis and haemostasis.* 2002; 87(3):366–73. PMID: 11916066.
24. Belvini D, Salviato R, Radossi P, Pierobon F, Mori P, Castaldo G, et al. Molecular genotyping of the Italian cohort of patients with hemophilia B. *Haematologica.* 2005; 90(5):635–42. PMID: 15921378.
25. Dani C, Piechaczyk M, Audigier Y, El Sabouty S, Cathala G, Marty L, et al. Characterization of the transcription products of glyceraldehyde 3-phosphate-dehydrogenase gene in HeLa cells. *Eur J Biochem.* 1984; 145(2):299–304. <https://doi.org/10.1111/j.1432-1033.1984.tb08552.x> PMID: 6499844.
26. So BR, Di C, Cai Z, Venters CC, Guo J, Oh JM, et al. A Complex of U1 snRNP with Cleavage and Polyadenylation Factors Controls Telescripting, Regulating mRNA Transcription in Human Cells. *Mol Cell.* 2019; 76(4):590–9 e4. <https://doi.org/10.1016/j.molcel.2019.08.007> PMID: 31522989; PubMed Central PMCID: PMC6874754.
27. Paz I, Kostli I, Ares M Jr., Cline M, Mandel-Gutfreund Y. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Res.* 2014; 42(Web Server issue):W361–7. <https://doi.org/10.1093/nar/gku406> PMID: 24829458; PubMed Central PMCID: PMC4086114.
28. Machado-Neto JA, Lazarini M, Favaro P, Franchi GC Jr., Nowill AE, Saad ST, et al. ANKHD1, a novel component of the Hippo signaling pathway, promotes YAP1 activation and cell cycle progression in prostate cancer cells. *Exp Cell Res.* 2014; 324(2):137–45. <https://doi.org/10.1016/j.yexcr.2014.04.004> PMID: 24726915
29. Zhou Z, Jiang H, Tu K, Yu W, Zhang J, Hu Z, et al. ANKHD1 is required for SMYD3 to promote tumor metastasis in hepatocellular carcinoma. *Journal of experimental & clinical cancer research: CR.* 2019; 38(1):18. <https://doi.org/10.1186/s13046-018-1011-0> PMID: 30646949; PubMed Central PMCID: PMC6332640.
30. Dowdy SF. Overcoming cellular barriers for RNA therapeutics. *Nat Biotechnol.* 2017; 35(3):222–9. <https://doi.org/10.1038/nbt.3802> PMID: 28244992.
31. van Ommen GJ, van Deutekom J, Aartsma-Rus A. The therapeutic potential of antisense-mediated exon skipping. *Curr Opin Mol Ther.* 2008; 10(2):140–9. PMID: 18386226.
32. Voit T, Topaloglu H, Straub V, Muntoni F, Deconinck N, Campion G, et al. Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. *The Lancet Neurology.* 2014; 13(10):987–96. [https://doi.org/10.1016/S1474-4422\(14\)70195-4](https://doi.org/10.1016/S1474-4422(14)70195-4) PMID: 25209738.

33. Liu S, Asparuhova M, Brondani V, Ziekau I, Klimkait T, Schumperli D. Inhibition of HIV-1 multiplication by antisense U7 snRNAs and siRNAs targeting cyclophilin A. *Nucleic Acids Res.* 2004; 32(12):3752–9. <https://doi.org/10.1093/nar/gkh715> PMID: 15254276; PubMed Central PMCID: PMC484190.
34. Dominski Z, Marzluff WF. Formation of the 3' end of histone mRNA: getting closer to the end. *Gene.* 2007; 396(2):373–90. <https://doi.org/10.1016/j.gene.2007.04.021> PMID: 17531405.
35. Goyenvalle A, Babbs A, Wright J, Wilkins V, Powell D, Garcia L, et al. Rescue of severely affected dystrophin/utrophin-deficient mice through scAAV-U7snRNA-mediated exon skipping. *Human molecular genetics.* 2012; 21(11):2559–71. <https://doi.org/10.1093/hmg/dds082> PMID: 22388933; PubMed Central PMCID: PMC3349427.
36. Imbert M, Dias-Florencio G, Goyenvalle A. Viral Vector-Mediated Antisense Therapy for Genetic Diseases. *Genes.* 2017; 8(2). <https://doi.org/10.3390/genes8020051> PMID: 28134780; PubMed Central PMCID: PMC5333040.
37. Oh JM, Di C, Venters CC, Guo J, Arai C, So BR, et al. U1 snRNP telescripting regulates a size-function-stratified human genome. *Nat Struct Mol Biol.* 2017; 24(11):993–9. <https://doi.org/10.1038/nsmb.3473> PMID: 28967884; PubMed Central PMCID: PMC5685549.
38. Jirka SM, Tanganyika-de Winter CL, Boertje-van der Meulen JW, van Putten M, Hiller M, Vermue R, et al. Evaluation of 2'-Deoxy-2'-fluoro Antisense Oligonucleotides for Exon Skipping in Duchenne Muscular Dystrophy. *Molecular therapy Nucleic acids.* 2015; 4:e265. <https://doi.org/10.1038/mtna.2015.39> PMID: 26623937; PubMed Central PMCID: PMC5014533.
39. Chi X, Gatti P, Papoian T. Safety of antisense oligonucleotide and siRNA-based therapeutics. *Drug discovery today.* 2017; 22(5):823–33. <https://doi.org/10.1016/j.drudis.2017.01.013> PMID: 28159625.
40. Eckenfelder A, Tordo J, Babbs A, Davies KE, Goyenvalle A, Danos O. The Cellular Processing Capacity Limits the Amounts of Chimeric U7 snRNA Available for Antisense Delivery. *Molecular therapy Nucleic acids.* 2012; 1:e31. <https://doi.org/10.1038/mtna.2012.24> PMID: 23344083; PubMed Central PMCID: PMC3390224.
41. Nathwani AC, Reiss UM, Tuddenham EG, Rosales C, Chowdary P, McIntosh J, et al. Long-term safety and efficacy of factor IX gene therapy in hemophilia B. *N Engl J Med.* 2014; 371(21):1994–2004. <https://doi.org/10.1056/NEJMoa1407309> PMID: 25409372; PubMed Central PMCID: PMC4278802.
42. Goyenvalle A, Wright J, Babbs A, Wilkins V, Garcia L, Davies KE. Engineering multiple U7snRNA constructs to induce single and multiexon-skipping for Duchenne muscular dystrophy. *Mol Ther.* 2012; 20(6):1212–21. Epub 2012/02/23. <https://doi.org/10.1038/mt.2012.26> mt201226 [pii]. PMID: 22354379; PubMed Central PMCID: PMC3369406.
43. Weidenfeld I, Gossen M, Low R, Kentner D, Berger S, Gorlich D, et al. Inducible expression of coding and inhibitory RNAs from retargetable genomic loci. *Nucleic Acids Res.* 2009; 37(7):e50. <https://doi.org/10.1093/nar/gkp108> PMID: 19264799; PubMed Central PMCID: PMC2673444.
44. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* 2015; 43(Database issue):D146–52. <https://doi.org/10.1093/nar/gku1104> PMID: 25378301; PubMed Central PMCID: PMC4383922.
45. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *Journal of computational biology: a journal of computational molecular cell biology.* 1997; 4(3):311–23. <https://doi.org/10.1089/cmb.1997.4.311> PMID: 9278062.