



**HAL**  
open science

# Détermination des mécanismes physiopathologiques d'anomalies rares de la coagulation à l'aide de modèles *in vitro* et d'approches génétiques innovantes

Yohann Jourdy

## ► To cite this version:

Yohann Jourdy. Détermination des mécanismes physiopathologiques d'anomalies rares de la coagulation à l'aide de modèles *in vitro* et d'approches génétiques innovantes. Médecine humaine et pathologie. Université de Lyon, 2017. Français. NNT : 2017LYSE1316 . tel-01696105

**HAL Id: tel-01696105**

**<https://theses.hal.science/tel-01696105>**

Submitted on 30 Jan 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT :

**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N° 205**  
**(Ecole Doctorale Interdisciplinaire Science Santé EDISS)**

**Spécialité de doctorat : Biologie médicale**  
**Discipline : Hémostasie**

Soutenue publiquement 20/12/2017, par :  
**Yohann Jourdy**

---

**Détermination des mécanismes  
physiopathologiques d'anomalies rares  
de la coagulation à l'aide de modèles *in  
vitro* et d'approches génétiques  
innovantes**

---

Devant le jury composé de :

Pr Négrier Claude	PU-PH Université de Lyon 1	Président
Pr Morange Pierre	PU-PH Université d'Aix-Marseille	Rapporteur
Pr Nguyen Philippe	PU-PH Université de Reims	Rapporteur
Pr Hermans Cédric	PU-PH CHU de Saint Luc	Examineur
Pr Vinciguerra Christine	PU-PH Université de Lyon 1	Directrice de thèse
Pr Dargaud Yesim	PU-PH Université de Lyon 1	Co-directrice de thèse

# UNIVERSITE CLAUDE BERNARD - LYON 1

## **Président de l'Université**

**M. le Professeur Frédéric FLEURY**

Président du Conseil Académique

M. le Professeur Hamda BEN HADID

Vice-président du Conseil d'Administration

M. le Professeur Didier REVEL

Vice-président du Conseil Formation et Vie Universitaire

M. le Professeur Philippe CHEVALIER

Vice-président de la Commission Recherche

M. Fabrice VALLÉE

Directrice Générale des Services

Mme Dominique MARCHAND

## ***COMPOSANTES SANTE***

Faculté de Médecine Lyon Est – Claude Bernard

Directeur : M. le Professeur G.RODE

Faculté de Médecine et de Maïeutique Lyon Sud – Charles  
Mérieux

Directeur : Mme la Professeure C. BURILLON

Faculté d'Odontologie

Directeur : M. le Professeur D. BOURGEOIS

Institut des Sciences Pharmaceutiques et Biologiques

Directeur : Mme la Professeure C. VINCIGUERRA

Institut des Sciences et Techniques de la Réadaptation

Directeur : M. X. PERROT

Département de formation et Centre de Recherche en Biologie  
Humaine

Directeur : Mme la Professeure A-M. SCHOTT

# ***COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE***

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

## **Remerciements**

Mes travaux ont été réalisés dans l'unité de recherche EA4609 de l'université Claude Bernard Lyon 1 en étroite collaboration avec le laboratoire de diagnostic génétique des maladies hémorragiques des Hospices Civils de Lyon. Ce travail a été réalisé sous la direction de la Professeure Christine Vinciguerra et le co-encadrement de la professeure Yesim Dargaud. Une partie de ce travail a été effectuée en collaboration avec le laboratoire de cytogénétique constitutionnelle des Hospices Civils de Lyon.

### **A madame la professeure Christine Vinciguerra,**

Je tiens à vous témoigner particulièrement toute ma gratitude pour la confiance que vous m'avez accordée dès mon arrivée au laboratoire en 2012 en acceptant de m'encadrer successivement pour mon Master 2, mon mémoire de DES et maintenant ma thèse d'université. Un grand merci pour la qualité de votre encadrement, votre soutien sans faille, vos encouragements en toutes circonstances et surtout pour votre grande disponibilité. Enfin, je vous remercie pour vos conseils avisés et pour tout l'intérêt que vous portez à mon avenir professionnel.

### **A madame la professeure Yesim Dargaud,**

Vous avez eu la tâche d'assurer mon co-encadrement durant cette thèse et je vous en suis grandement reconnaissant. Travailler à vos côtés m'a été extrêmement bénéfique, et notamment sur le plan scientifique. Vous m'avez montré ce qu'était la persévérance en recherche. C'est une grande chance pour moi de pouvoir continuer à collaborer avec vous sur d'autres projets au sein de l'EA 4609.

### **A monsieur le Professeur Claude Négrier,**

Je vous remercie pour m'avoir accordé votre confiance toutes ces années et de m'avoir accueillie à la fois dans votre service et dans votre équipe de recherche. J'espère que ce travail de thèse sera à la hauteur de vos attentes.

**A Messieurs les Professeurs Pierre Morange et Philippe Nguyen,**

J'ai eu le plaisir de faire vos connaissances lors du dernier congrès du GFHT-CoMETH. Je souhaite vous exprimer toute ma gratitude pour avoir accepté de consacrer une partie de votre temps précieux à l'analyse de mon travail et d'en être les rapporteurs.

**A monsieur le Professeur Cédric Hermans,**

J'ai aussi eu l'honneur de vous rencontrer lors du dernier congrès du GFHT-CoMETH. Je vous adresse tous mes remerciements pour avoir accepté de lire ce manuscrit et de faire partie de ce Jury.

\*\*\*\*\*

Je tiens également à témoigner ma plus grande reconnaissance à toutes les personnes ayant contribué à la réalisation de ce travail :

Mathilde Fretigny pour ses conseils avisés sur l'interprétation des variants en génétique humaine.

Le professeur Damien Sanlaville et Nicolas Chatron pour votre collaboration dans l'étude des grands réarrangements des gènes *F8* et *F9*.

Dominique Bozon pour votre aide précieuse et l'apport de vos compétences sur les anomalies d'épissage.

Nathalie pour tes précieux conseils en culture cellulaire et ta collaboration sur le projet thrombomoduline.

Alexandre pour ta collaboration à une partie de ce travail, tes conseils sur l'utilisation des siRNA et ton aide technique qui a été décisive.

Aux membres du laboratoire de diagnostic génétique des maladies hémorragiques et aux membres de l'EA 4609 pour toute l'aide fournie durant la réalisation de ces travaux et plus particulièrement à Dorothée et Stéphanie.

\*\*\*\*\*

Un grand merci aussi à toutes les personnes qui contribuent tous les jours à rendre le travail au laboratoire agréable et stimulant, c'est-à-dire ... tout le monde (biologistes, techniciens et secrétaires) !!! Plus particulièrement à :

Christophe, tu as été mon mentor lors de mes premiers pas en hémostase, ton dynamisme au laboratoire rend le travail à tes côtés très motivant.

Mes co-assistantes en hématologie, Fanélie et Emilie, pour les séances de radio potin.

A Sandra, tu y es presque !!! J'ai hâte de venir assister prochainement à ta soutenance de thèse.

\*\*\*\*\*

Je souhaite remercier toutes ma famille pour leur soutien de tous les jours et leurs encouragements, mes parents, mes sœurs Sarah et Marion, mon beau-frère Guillaume ainsi que les deux petites princesses Chloé et Emma. Et non Chloé, je ne serais toujours pas habillé en pirate aujourd'hui !

Je remercie aussi ma belle-famille qui m'a toujours accueilli chaleureusement bien que je leur aie dérobé un de leur diamant cet été.

Enfin, je tiens à exprimer toute ma reconnaissance et tout mon amour à Marion qui a connu ces dernières années tous les coulisses de cette thèse. Merci pour ton soutien sans faille, pour ta bonne humeur communicative, pour la confiance que tu m'as toujours voué et surtout pour le réconfort que tu as toujours su me donner dans les moments difficiles. Déjà 9 ans de bonheur... et le meilleur reste à venir.

# Table des matières

<b>REMERCIEMENTS.....</b>	<b>4</b>
<b>LISTE DES FIGURES .....</b>	<b>10</b>
<b>LISTE DES ABREVIATIONS.....</b>	<b>13</b>
<b>INTRODUCTION.....</b>	<b>15</b>
<b>SECTION 1 : GENETIQUE DES DEFICITS HEREDITAIRES DE LA COAGULATION .....</b>	<b>18</b>
<b>Partie 1 : Principes généraux du système de la coagulation .....</b>	<b>19</b>
<b>Partie 2 : Les déficits héréditaires en facteur de coagulation.....</b>	<b>22</b>
1. Les hémophilies .....	22
2. Les déficits rares de la coagulation .....	26
<b>Partie 3 : Génétique des déficits héréditaires en facteur de la coagulation : exemple de l'hémophilie A. ....</b>	<b>29</b>
1. Structure du gène <i>F8</i> .....	29
2. Origine des mutations du gène <i>F8</i> .....	32
3. Les mutations du gène <i>F8</i> .....	33
<b>Partie 4: Maladies hémorragiques héréditaires liées à des mutations gain de fonction. ....</b>	<b>48</b>
1. Mutations gain de fonction localisées dans le gène <i>F5</i> : FV Texas et FV Amsterdam.....	48
2. Mutations gain de fonction localisées dans le gène <i>THBD</i> : c.1611C>A (p.Cys537X). ....	49
<b>SECTION 2. ANOMALIES GENETIQUES DE GRANDE TAILLE DANS L'HEMOPHILIE:</b>	
<b>MISE EN EVIDENCE ET INTERET POUR LE CONSEIL GENETIQUE. ....</b>	<b>51</b>
<b>Partie 1. Caractérisation par analyse chromosomique sur puce à ADN de 6 délétions totales du gène <i>F9</i>:</b>	
<b>implication de la délétion du gène <i>SOX3</i> dans les déficits mentaux observés chez certains patients.....</b>	<b>52</b>



3. Résumé .....	53
4. Introduction .....	54
5. Matériels et Méthodes .....	55
6. Résultats et discussion .....	57

**Partie 2. Caractérisation par analyse chromosomique sur puce à ADN de cinq réarrangements génomiques**

**complexes affectant le gène *F8* : intérêt pour le conseil génétique. .... 62**

1. Résumé .....	63
2. Introduction .....	64
3. Matériels et méthodes.....	64
4. Résultats.....	68
5. Discussion .....	76

**SECTION 3. CARACTERISATION DE VARIANTS DE SIGNIFICATION INDETERMINEE**

**LOCALISES AU NIVEAU DES SITES D'ÉPISSAGE DU GÈNE *F8*. .... 81**

**Partie 1 : Rappels bibliographiques sur la réaction d'épissage..... 82**

1. La réaction d'épissage.....	82
2. Le cycle d'assemblage du spliceosome .....	83
3. Éléments participant à la reconnaissance de l'exon par le splicéosome .....	87

**Partie 2. Etude *in vitro* de l'impact sur l'épissage de 26 variations du gène *F8* ..... 95**

1. Résumé .....	95
2. Introduction .....	96
3. Matériels et méthodes.....	97
4. Résultats.....	112
5. Discussion .....	120
6. Annexes.....	127

## **SECTION 4 : DESCRIPTION DE NOUVEAUX MECANISMES MOLECULAIRES**

### **RESPONSABLES DE PATHOLOGIES HEMORRAGIQUES HEREDITAIRES..... 129**

#### **Partie 1 : Exonisation d'une séquence *AluY* liée à une anomalie de fixation de l'hnRNP C: un mécanisme**

##### **original responsable d'hémophilie A mineure ..... 130**

1. Résumé ..... 131
2. Rappels bibliographiques : les séquences *Alu*..... 133
3. Introduction ..... 140
4. Matériels et méthodes..... 141
5. Résultats..... 148
6. Discussion ..... 153
7. Annexes..... 156

#### **Partie2 : Etude des effets fonctionnels de la mutation c.1611C>A (p.Cys537X) du gène *THBD*. ..... 158**

1. Résumé ..... 159
1. Rappels bibliographiques sur la thrombomoduline ..... 161
2. Introduction ..... 166
3. Matériels et méthodes..... 166
4. Résultats..... 169
5. Discussion ..... 176

### **CONCLUSIONS ET PERSPECTIVES..... 180**

### **ANNEXES ..... 184**

1. Article en revision (Américan Journal of Human Genetic): Recurrent *F8* intronic deletion in mild haemophilia A causes *Alu* exonization due to the decrease of hnRNP C binding. .... 184

### **BIBLIOGRAPHIE ..... 185**

## Liste des figures

FIGURE 1 : SCHEMA SIMPLIFIE DE LA COAGULATION .....	20
FIGURE 2 : STRUCTURE DU GENE <i>F8</i> .....	31
FIGURE 3: STRUCTURE DU FVIII INTRACELLULAIRE, CIRUCLANT ET DU FVIII ACTIVE. ....	32
FIGURE 4 : REARRANGEMENT GENOMIQUE RESULTANT D'UNE RECOMBINAISON HOMOLOGUE NON ALLELIQUE ENTRE DEUX SEQUENCES REPETEES LCRS.....	34
FIGURE 5 : L'INVERSION DE L'INTRON 22 DE TYPE 1 (65) .....	35
FIGURE 6 : JONCTION D'EXTREMITE NON-HOMOLOGUE NHEJ .....	37
FIGURE 7 : JONCTION D'EXTREMITE PRESENTANT DES MICROHOMOLOGIES (MMEJ). ....	37
FIGURE 8 : FORK STALLING AND TEMPLATE SWITCHING (FOSTES). ....	38
FIGURE 9 : REPARATION DE L'ADN PAR EXCISION DE BASE (BER) .....	41
FIGURE 10 : REPARATION DE L'ADN PAR EXCISION DE NUCLEOTIDE (NER) .....	42
FIGURE 11 : LOCALISATION DES MUTATIONS FAUX-SENS SUR LE FACTEUR VIII .....	44
FIGURE 12 : COORDONNEES GENOMIQUES DES DIFFERENTES DELETIONS ETUDIEES.....	59
FIGURE 13 : PROFIL ANORMAL DE L'INVERSION DE L'INTRON 1 .....	69
FIGURE 14 : COORDONNEES GENOMIQUES DES DIFFERENTS REARRANGEMENTS ETUDIES.....	70
FIGURE 15 : RESULTATS DE SOUTHERN BLOT POUR LE PATIENT 2.....	71
FIGURE 16 : RESULTATS DE SOUTHERN BLOT POUR LE PATIENT 3.....	72
FIGURE 17 : RESULTATS DE SOUTHERN BLOT POUR LE PATIENT 4.....	73
FIGURE 18 : RESULTATS DE SOUTHERN BLOT POUR LA PATIENTE 5 .....	74
FIGURE 19 : STRUCTURE DES JONCTIONS INTRON/EXON.....	82
FIGURE 20 : LA REACTION D'EPISSAGE.....	83
FIGURE 21 : COMPOSITION DES PRINCIPAUX SNRNPS COMPOSANT LE SPLICEOSOME CHEZ L'HOMME .....	84
FIGURE 22 : CYCLE D'ASSEMBLAGE DU SPLICEOSOME .....	85
FIGURE 23 : INTERACTIONS MOLECULAIRES AU NIVEAU DU SITE ACCEPTEUR D'EPISSAGE AU SEIN DU COMPLEXE E DU SPLICEOSOME .....	86
FIGURE 24 : INTERACTIONS MOLECULAIRES AU NIVEAU DU SITE ACCEPTEUR D'EPISSAGE AU SEIN DU COMPLEXE A DU SPLICEOSOME .....	86

FIGURE 25 : LA SEQUENCE CONSENSUS DU SITE DONNEUR D'ÉPISSAGE.....	88
FIGURE 26 : SCHEMATISATION DES DEUX MODELES DE RECONNAISSANCE DES SITES D'ÉPISSAGE PAR LE SPLICEOSOME.....	90
FIGURE 27 : LISTE ET STRUCTURE DES PROTEINES SR.....	92
FIGURE 28 : LES PROTEINES HNRNP.....	93
FIGURE 29 : REPRESENTATION SCHEMATIQUE DU MINIGENE UTILISE DANS CE TRAVAIL.....	100
FIGURE 30 : EXEMPLE DE MIGRATION DES PRODUITS DE PCR ISSUS DE LA VERIFICATION DES CLONES BACTERIENS TRANSFORMES AVEC LE MINIGENE CONTENANT L'EXON 5 DU GENE <i>F8</i> . ....	104
FIGURE 31 : EXEMPLE DE MIGRATION SUR GEL D'AGAROSE 1% DES ARNM TOTAUX PURIFIES EXTRAITS A PARTIR DE LYSATS DE CELLULES HELA.....	106
FIGURE 32 : EXEMPLE D'ANALYSE DES TRANSCRITS PRODUITS PAR DES CELLULES HELA APRES TRANSFECTION PAR LES MINIGENES <i>F8</i> EXON 5 WILD-TYPE (EXON NORMAL) ET <i>F8</i> EXON 5 C.670G>A.....	107
FIGURE 33 : EXEMPLE DE CLONAGE TOPO REALISE POUR L'ETUDE DE LA MUTATION C.6066C>T.....	108
FIGURE 34 : RESULTATS DE L'ANALYSE <i>IN VITRO</i> EFFECTUEE APRES TRANSFECTION DE CELLULES HELA POUR LES 18 VARIATIONS SITUEES AU NIVEAU D'UN SITE DONNEUR D'ÉPISSAGE. ....	115
FIGURE 35 : RESULTATS DE L'ANALYSE <i>IN VITRO</i> EFFECTUEE APRES TRANSFECTION DE CELLULES HUH7 POUR LES 18 VARIATIONS SITUEES AU NIVEAU D'UN SITE DONNEUR D'ÉPISSAGE.....	115
FIGURE 36 : ACTIVITE DU FVIII MESUREE PAR METHODE CHRONOMETRIQUE (FVIII:C) ET CHROMOGENIQUE (FVIII:CR) ET FVIII ANTIGENE (FVIII:AG) MESURE DANS LES SURNAGEANTS DE CULTURE CELLULAIRE.....	116
FIGURE 37 : RESULTATS DE L'ANALYSE <i>IN VITRO</i> EFFECTUEE APRES TRANSFECTION DE CELLULES HELA POUR LES 8 VARIATIONS SITUEES AU NIVEAU D'UN SITE ACCEPTEUR D'ÉPISSAGE.....	119
FIGURE 38 : RESULTATS DE L'ANALYSE <i>IN VITRO</i> EFFECTUEE APRES TRANSFECTION DE CELLULES HUH7 POUR LES 8 VARIATIONS SITUEES AU NIVEAU D'UN SITE ACCEPTEUR D'ÉPISSAGE.....	119
FIGURE 39 : STRUCTURE D'UNE SEQUENCE <i>ALU</i> .....	134
FIGURE 40 : MECANISME DE RETROTRANSPOSITION DES SEQUENCES <i>ALU</i> .....	135
FIGURE 41 : EXEMPLE D'AMPLIFICATION DES TRANSCRITS DU GENE <i>ACTB</i> .....	144
FIGURE 42 : ANALYSE DES TRANSCRITS DU GENE <i>F8</i> PAR MIGRATION SUR GEL D'AGAROSE 1%.....	145
FIGURE 43 : IDENTIFICATION DE LA DELETION INTRONIQUE C.2113+461_2113+473DEL CHEZ LES DEUX CAS INDEX.....	149

FIGURE 44 : RESULTATS DE L'ANALYSE FONCTIONELLE <i>IN VITRO</i> DE L'IMPACT DE LA DELETION C.2113+461_2113+473DEL SUR L'EXONISATION DE LA SEQUENCE <i>ALUY</i> .....	150
FIGURE 45 : L'INSERTION DE MUTATION DANS LA QUEUE POLY-(T) DE LA SEQUENCE <i>ALUY</i> FAVORISE SON EXONNISATION.....	151
FIGURE 46 : L'INHIBITION DE LA PRODUCTION D'HNRNP C FAVORISE L'EXONISATION DE LA SEQUENCE <i>ALUY</i> DANS LES CELLULES HELA.....	152
FIGURE 47 : STRUCTURE DE LA THROMBOMODULINE (276).....	162
FIGURE 48 : EXPRESSION CELLULAIRE DES TM-WT ET TM <sub>1-536</sub> (P .CYS537X).....	170
FIGURE 49 : ANALYSE PAR IMMUNOFLUORESCENCE DE LA THROMBOMODULINE A LA SURFACE DES CELLULES COS-1.....	171
FIGURE 50 : LIBERATION DE THROMBOMODULINE DANS LES SURNAGEANTS CELLULAIRES SOUS DIFFERENTES CONDITIONS .....	173
FIGURE 51 : RELATION ENTRE LA TAILLE DU DOMAINE TRANSMEMBRANAIRE ET LA LIBERATION DE THROMBOMODULINE DANS LES SURNAGEANTS CELLULAIRES.....	175

## Liste des abréviations

3'ss :	Site accepteur d'épissage
3'UTR :	région 3' non transcrite
5'ss :	Site donneur d'épissage
AA :	Acides aminés
ADAR :	Adenosine Deaminase Acting on RNA
BER :	Réparation par excision de base
BP :	Point de branchement
CAT :	Calibrated Automated Thrombography
CMA :	Analyse chromosomique sur puce à ADN
CNV :	Variation du nombre de copie
EGF :	Epidermal Growth Factor
EPCR :	Récepteur endothélial de la protéine C
ESE :	Séquence exonique activatrice
ESS :	Séquence intronique inhibitrice
ETP :	Potentiel endogène de thrombine
F :	Facteur
FT :	Facteur tissulaire
FII :	Prothrombine
FIIa :	Thrombine
FIX:C :	Activité coagulométrique du FIX
FIX:Cr :	Activité chromogénique du FIX
FoSTeS :	Fork stalling and template switching
FVIII:C :	Activité coagulométrique du FVIII
FVIII:Cr :	Activité chromogénique du FVIII
HA :	Hémophilie A
HB :	Hémophilie B
hnRNP :	Heterogeneous Nuclear Ribonucleoproteins
Inv1 :	Inversion de l'intron 1
Inv22 :	Inversion de l'intron 22
ISE :	Séquence intronique activatrice
ISS :	Séquence intronique inhibitrice
IV :	Voie intraveineuse
LINE :	Long interspersed element
LR-PCR :	PCR long-fragment
LTRs :	Long Terminal Repeats
MLPA :	Multiplex Ligation-Dependent Probe Amplification
MMBIR :	Microhomology-mediated break-induced replication (MMBIR)
MMEJ :	Jonction d'extrémité présentant des microhomologies
MMP :	Métalloprotéase matricielle
MMR :	DNA Mismatch Repair
NAHR :	Recombinaison homologue non allélique
NER :	Réparation par excision de nucléotides
NHEJ :	Jonction d'extrémité non-homologue
PAR-1 :	Récepteur plaquettaire de la thrombine
PC :	Protéine C
PF4 :	Facteur 4 plaquettaire
PPP :	Plasma pauvre en plaquettes
PS :	Protéine S
snRNA :	Petit ARN nucléaire
snRNP :	Petites ribonucleoparticules nucléaires
TAFI :	Inhibiteur de la fibrinolyse activable par la thrombine
TCA :	Temps de céphaline avec activateur

TGT : Test de génération de thrombine  
TFPI: Inhibiteur de la voie du facteur tissulaire  
SB : Southern Blot  
sTM: Thrombomoduline soluble  
SVF: Serum de veau foetal  
TM : Thrombomoduline  
VWF : Facteur von Willebrand

## Introduction

La coagulation est l'un des processus physiologiques participant au maintien de l'intégrité des vaisseaux sanguins. A la suite d'une lésion vasculaire, un ensemble de protéines, appelées facteurs de la coagulation, sont activées en cascade pour aboutir à la génération de thrombine. L'une des principales fonctions de la thrombine est de produire la fibrine, élément indispensable à la formation d'un caillot sanguin stable. Les déficits congénitaux en l'une ou plusieurs protéines de la cascade de la coagulation sont responsables de pathologies hémorragiques dont les plus fréquentes sont l'hémophilie A (déficit en Facteur VIII de la coagulation) et l'hémophilie B (déficit en Facteur IX de la coagulation).

Le développement des méthodes de biologie moléculaire à partir des années 1990 a permis l'étude des gènes codant les protéines de la coagulation. L'identification des anomalies génétiques affectant ces gènes a engendré à la fois une meilleure compréhension de la physiologie de la coagulation et une meilleure prise en charge des patients. En effet, la connaissance de la mutation responsable chez un patient permet l'identification avec certitude des individus pouvant transmettre la pathologie dans sa famille. Un conseil génétique peut ainsi être proposé à la famille, aux couples ou aux individus porteurs de la mutation afin de les renseigner au mieux sur leur pathologie mais aussi sur les risques possibles pour leur descendance.

L'apparition des méthodes de séquençage à haut débit a entraîné une augmentation importante du nombre de variations génétiques identifiées dans les gènes associés aux pathologies congénitales hémorragiques. Cependant, la majorité des variants identifiés sont spécifiques d'une famille et la détermination du caractère pathogène ou non de ces variants est difficile. Ces anomalies génétiques sont alors considérées comme des variations de signification indéterminée. Le développement de méthodes *in vitro* permettant de démontrer le caractère pathogène de ces anomalies est alors d'une grande aide pour le conseil génétique.



En parallèle, le développement de techniques d'étude du génome dans sa globalité telle que l'analyse chromosomique sur puce à ADN (CMA) a permis la mise en évidence d'anomalies génétiques de grande taille (délétions et duplications) ne se limitant pas aux seuls gènes codant des protéines de la coagulation mais affectant aussi les gènes localisés dans leur environnement. Il est alors nécessaire, en vue du conseil génétique, de rechercher les conséquences possibles sur le phénotype des patients du dysfonctionnement de chacun de ces gènes.

Ce travail de thèse comporte trois parties. La première partie a consisté en la caractérisation par CMA de réarrangements génomiques de grande taille mis en évidence chez des patients hémophiles A ou B par le laboratoire de biologie moléculaire du Centre Hospitalier Universitaire de Lyon. La caractérisation précise de ces anomalies avait pour objectif l'étude des corrélations génotypes/phénotypes chez ces patients.

La deuxième partie de cette thèse a été concentrée sur l'étude de variants de signification indéterminée localisés au niveau des sites d'épissage du gène *F8*. L'objectif de ce travail était de développer une technique *in vitro* de type minigène afin de démontrer la pathogénicité de ces variations génétiques.

La troisième partie a concerné l'étude de patients pour lesquels aucune anomalie génétique n'a été détectée dans un premier temps. Dans cette situation, l'identification des membres de la famille pouvant transmettre l'anomalie est délicate. La mise en place de techniques de biologie moléculaire non utilisées en routine telles que les analyses de transcrits et les analyses d'exome peut permettre la découverte de nouvelles anomalies génétiques. Le lien de cause à effet entre l'anomalie génétique et le phénotype du patient doit ensuite être démontré. Cette partie comporte deux études. Dans la première, nous rapportons une délétion intronique profonde retrouvée chez plus de la moitié des hémophiles A mineurs de la cohorte lyonnaise chez lesquels aucune anomalie génétique n'avait été détectée à ce jour. Nous décrivons un mécanisme moléculaire original par lequel cette délétion est responsable de l'exonisation partielle d'une séquence *Alu*. Dans une seconde étude, nous avons étudié la mutation c.1611C>A du gène *THBD* codant la thrombomoduline. Cette anomalie a été identifiée lors d'une

analyse d'exome par l'équipe de Lyon chez une patiente présentant un syndrome hémorragique grave inexpliqué. L'objectif de cette dernière étude a été d'explorer les mécanismes physiopathologiques induits par la mutation c.1611C>A permettant d'expliquer les taux très élevés de thrombomoduline soluble observés chez cette patiente.

Après un chapitre dédié à des rappels bibliographiques, ces trois parties seront développées sur la base de publications acceptées ou en cours de soumission.

## **Section 1 : Génétique des déficits héréditaires de la coagulation**

## Partie 1 : Principes généraux du système de la coagulation

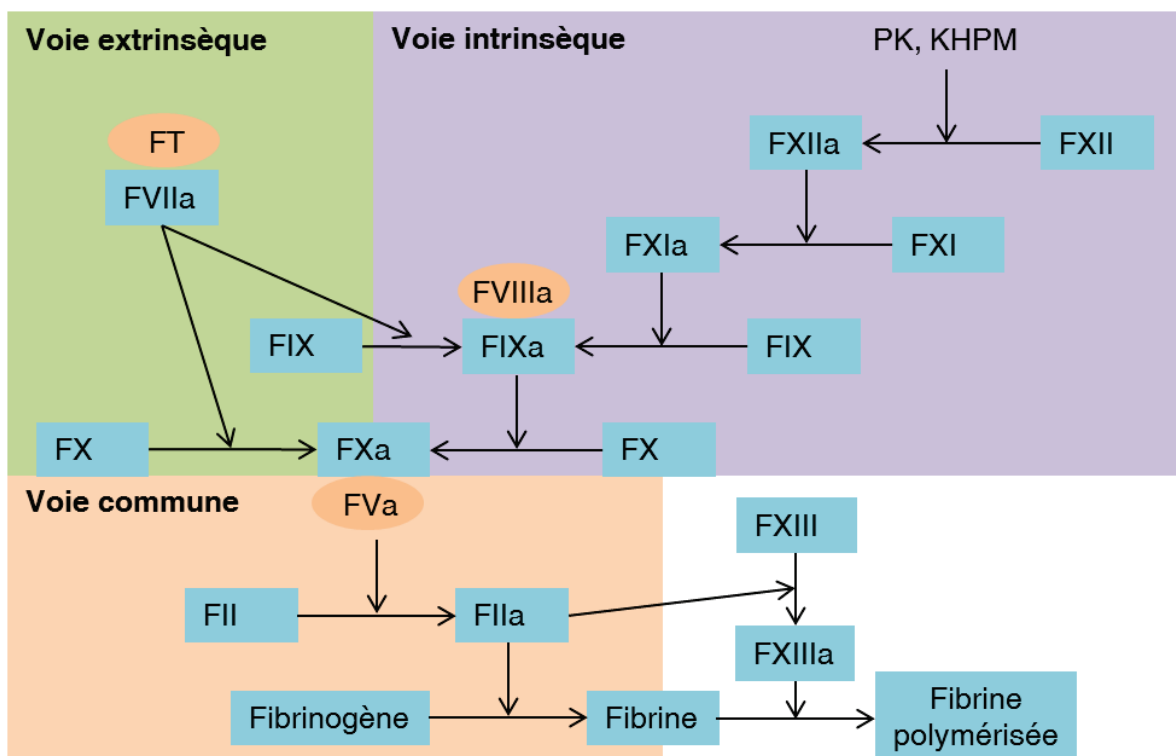
L'hémostase est un processus physiologique permettant de restaurer l'intégrité des vaisseaux sanguins suite à une brèche vasculaire. Classiquement, l'hémostase sanguine est divisée en trois étapes successives : l'hémostase primaire, la coagulation et la fibrinolyse. Lors de l'hémostase primaire, les plaquettes et le facteur Von Willebrand (VWF) forment très rapidement un thrombus blanc afin de limiter la perte sanguine. Ce thrombus blanc est consolidé lors de la coagulation par un réseau de fibrine insoluble. Plus tardivement, après réparation du vaisseau, le caillot est dégradé lors de la fibrinolyse.

De nombreux facteurs d'origine cellulaire (plaquettes activées à la surface desquelles a lieu la réaction de coagulation, cellules exprimant le facteur tissulaire (FT)) ou protéique (facteurs (F) de coagulation, cofacteurs et fibrinogène) interviennent lors du processus de coagulation. La majorité des facteurs de coagulation sont des enzymes appartenant soit à la famille des sérines protéases (FII, FVII, FIX, FX, FXI, FXII et kallikréine) soit à la famille des transglutaminases (FXIII). Ces facteurs sont sécrétés par le foie sous forme de zymogènes inactifs (pro-facteurs) qui deviendront actifs après clivage protéolytique. La production des FII, FVII, FX et FIX par le foie est dépendante de la présence de vitamine K. L'activité de plusieurs de ces facteurs est potentialisée grâce à l'action de cofacteurs protéiques. Les FV, FVIII et le kininogène de haut poids moléculaire (KHPM) sont respectivement les cofacteurs des FX activé (FXa), FIX activé (FIXa) et de la kallikréine.

Le processus de coagulation peut être divisé en plusieurs étapes. La première est la phase d'initiation qui se déclenche dès l'apparition d'une lésion vasculaire. Cette phase correspond à la voie extrinsèque de la coagulation (Figure 1). Suite à une brèche vasculaire, le sous endothélium est en contact direct avec la circulation sanguine. Les cellules formant ce tissu expriment fortement le FT à leur surface. L'initiation de la cascade de coagulation est déclenchée par la liaison du FVII activé (FVIIa) circulant au FT. Le complexe FT/FVIIa catalyse l'activation des FX et FIX en FXa et FIXa. La faible quantité de FXa produite lors de cette étape permet l'activation d'une petite fraction de la prothrombine (FII) en thrombine (FIIa).

La seconde phase de la coagulation correspond à la phase d'amplification (Figure 1). Durant cette étape, la thrombine générée à l'état de trace au niveau de la lésion vasculaire amplifie sa propre production. Ainsi, le FIIa active le FXI en FXIa et les cofacteurs VIII et V en FVIIIa et FVa activé (FVa). Le FVIIIa s'associe avec le FIXa à la surface des plaquettes activées pour former le complexe Tenase. Ce complexe a la capacité de générer du FXa beaucoup plus efficacement que le complexe FT/FVIIa. De la même manière, le FXa produit s'associe au FVa à la surface des plaquettes activées pour former le complexe prothrombinase qui génère de la thrombine beaucoup plus efficacement que le FXa seul. Cette étape de la coagulation est ainsi caractérisée par une explosion de la génération de thrombine.

La dernière étape de la coagulation est l'étape de fibrinoformation. La très grande quantité de thrombine générée lors de la phase d'amplification transforme le fibrinogène en monomère de fibrine soluble et active le FXIII en FXIIIa. Les filaments de fibrine se polymérisent au niveau du caillot. Ce caillot est ensuite consolidé par l'action du FXIIIa qui catalyse la formation de liaisons covalentes entre les fibres de fibrines afin de former un réseau de fibrine insoluble.



**Figure 1 : Schéma simplifié de la coagulation.** PK, prékallitréine ; KHPM, kininogène de haut poids moléculaire.

Afin que la réaction de coagulation ne se propage pas dans les vaisseaux non lésés, de nombreux systèmes inhibiteurs permettent de contrôler ce processus. Les principaux systèmes inhibiteurs de la coagulation sont :

- l'inhibiteur de la voie du facteur tissulaire (TFPI) qui régule négativement l'activité du complexe FT/FVIIa

- l'antithrombine qui est un inhibiteur direct des sérines protéases (FIIa, FXa, FXIa essentiellement)

- le système protéine C (PC) / protéine S (PS) qui inhibe la boucle d'amplification de la génération de thrombine en clivant les FVa et FVIIIa. Dans ce système, la thrombine se lie à la thrombomoduline (TM), une protéine transmembranaire exprimée à la surface des cellules endothéliales. Au sein du complexe TM/FIIa ; la thrombine perd ses propriétés procoagulantes et devient anticoagulante en catalysant l'activation de la PC en PC activée (PCa). Le rôle de la TM est développé plus en détail dans la section 4.

## Partie 2 : Les déficits héréditaires en facteur de coagulation

Les déficits en facteur de la coagulation forment un ensemble de pathologies hémorragiques héréditaires. La physiopathologie de ces maladies est commune, le déficit en l'un des facteurs de coagulation est responsable d'une diminution de la génération de thrombine entraînant la formation d'un caillot instable et de mauvaise qualité ne permettant pas de combler efficacement une brèche vasculaire. Les déficits en facteur de la coagulation peuvent avoir des expressions cliniques variables allant du syndrome hémorragique sévère à des formes pauci-symptomatiques. Les déficits en FVIII et FIX, appelés respectivement hémophilie A (HA, OMIM 306700) et B (HB, OMIM 306900), sont les plus fréquents. Les autres déficits en facteurs de coagulation, moins fréquents, sont rassemblés sous le nom générique de déficits rares de la coagulation.

### 1. Les hémophilies

Les hémophilies sont des pathologies hémorragiques congénitales dues à un déficit partiel ou complet en FVIII (HA) ou FIX (HB). Ces pathologies sont causées par des anomalies génétiques touchant les gènes codant le FVIII (*F8*) ou le FIX (*F9*), tous deux localisés sur le chromosome X. Leur transmission se fait sur le mode récessif lié à l'X. Leur incidence dans la population mondiale est estimée à 1/5000 et 1/30 000 naissances masculines respectivement pour l'HA et l'HB (1). En France, on dénombre environ 6000 hémophiles A et 1300 hémophiles B (<https://www.francecoag.org>, Octobre 2017). En fonction de l'activité résiduelle plasmatique des FVIII (FVIII:C) et FIX (FIX:C), les hémophiles sont classés en trois grades de sévérité : les hémophiles sévères (FVIII:C ou FIX:C <1%), modérés ( $1\% \leq$  FVIII:C ou FIX:C <5%) et mineurs ( $5\% \leq$  FVIII:C ou FIX:C <40%) (2). La répartition des hémophiles dans les différents grades de sévérité est présentée dans le Tableau 1.

Phénotype	Sévère, %	Modéré, %	Mineure, %
Hémophilie A	31	14	55
Hémophilie B	27	28	45

**Tableau 1 : Répartition des hémophiles dans les différents grades de sévérité en France.** D'après les données du réseau Francecoag. (<https://www.francecoag.org>, Octobre 2017).

## 1. Présentation clinique des hémophilies

La symptomatologie clinique des hémophilies est relativement bien corrélée aux taux de FVIII:C et FIX:C. Les hémophiles mineurs sont classiquement asymptomatiques, des hémorragies peuvent apparaître lors de traumatismes importants ou de procédures invasives telles que des chirurgies. Les hémophiles modérés sont caractérisés par l'apparition d'hémorragies suite à des traumatismes mineurs. Enfin, les hémophiles sévères présentent une symptomatologie hémorragique importante avec des saignements fréquents et spontanés essentiellement au niveau des articulations (hémarthroses) et des muscles (hématomes) mais aussi plus rarement au niveau d'autres localisations (hémorragies digestives, hématuries, hémorragies cérébrales) (3,4). La répétition des hémarthroses au niveau de la même articulation entraîne sa dégradation progressive et est responsable à terme d'une limitation fonctionnelle importante qui nécessite la pose d'une prothèse (arthropathie hémophilique).

## 2. Diagnostic biologique des hémophilies

Les hémophilies sont recherchées dans trois contextes principaux :

- Existence d'antécédents familiaux d'hémophilie
- Accident hémorragique inexpliqué
- Découverte fortuite sur un bilan de coagulation de routine

Sur le bilan de coagulation standard, une hémophilie doit être évoquée devant un allongement isolé du temps de céphaline avec activateur (ratio  $TCA_{\text{patient}}/TCA_{\text{témoin}} > 1,2$ ). Des dosages spécifiques de l'activité plasmatique des facteurs de la voie intrinsèque (FVIII, FIX et FXI) permettent de poser le diagnostic.

Il existe deux méthodes principales pour mesurer les activités plasmatiques des FVIII et FIX. La première est le dosage coagulométrique (FVIII:C et FIX:C). Il est basé sur le principe d'un TCA. Le plasma dilué du patient est mélangé à un plasma témoin déficitaire en FVIII ou FIX. Le temps de coagulation de ce mélange est inversement proportionnel à l'activité des FVIII ou FIX. La deuxième méthode est le dosage chromogénique (FVIII:Cr et FIX:Cr). Ce dosage est basé sur la génération de FXa par le complexe FVIIIa/FIXa. L'activité du FXa est mesurée à l'aide d'un substrat chromogène



spécifique. Les taux de FVIII et FIX sont directement proportionnels à l'activité du FXa générée. Les techniques chromogéniques et chromométriques sont équivalentes dans la majorité des cas. Cependant, une discordance entre les dosages FVIII:C et FVIII:Cr existe chez environ un tiers des patients hémophiles A mineurs (5).

Les principaux diagnostics différentiels de l'hémophilie A sont la maladie de Willebrand, le déficit combiné en FV+FVIII et l'hémophilie A acquise.

### 3. Traitement des hémophilies

La prévention ou le traitement des événements hémorragiques liés à l'hémophilie repose sur l'administration d'un traitement substitutif (concentrés de FVIII ou de FIX d'origine plasmatique ou recombinante) par voie intraveineuse (IV). Les traitements substitutifs peuvent être administrés soit à la demande lors d'un épisode hémorragique soit de manière prophylactique pour les hémophiles sévères. L'objectif du traitement prophylactique est de maintenir le taux de facteur résiduel au-dessus de 1-3% afin de prévenir les hémorragies spontanées (6). Le traitement prophylactique est actuellement considéré comme la référence dans le traitement des hémophilies sévères (7). Du fait des demi-vies des FVIII (8-12h) et FIX (18-24h), le traitement prophylactique des patients hémophiles sévères consiste en environ trois injections IV par semaine pour l'HA et deux injections IV par semaine pour l'HB.

Le nombre important d'injections est l'un des principaux facteurs de non observance chez les hémophiles, en particulier chez les enfants et les adolescents. De nombreuses molécules à demi-vie prolongée sont actuellement en développement (ou récemment arrivées sur le marché). Les stratégies utilisées pour augmenter la demi-vie des FVIII et FIX consistent soit en la fixation de groupement polyéthylène glycol (PEG) sur les molécules actives soit en la création de protéines de fusion (avec le fragment Fc des immunoglobulines ou avec l'albumine ; Tableau 2). Dans le cas de l'HB, ces nouvelles molécules ont permis une augmentation significative de la demi-vie du FIX (environ x 2-5) permettant un espacement des injections. L'augmentation de la demi-vie du FVIII est plus modeste (environ x 1,4-1,7) (8). Une dernière stratégie a été utilisée pour augmenter la demi-vie du FVIII. Elle

consiste en la production d'une protéine de FVIII composée d'une seule chaîne polypeptidique ce qui permet d'augmenter sa stabilité (9).

Pathologie	Molécule	Technologie	Demi-vie	Stade
Hémophilie A	BAY94-9027	Pégylation		Phase III
	N8-GP	Glycopégylation		Phase III
	Adynovate (BAX855)	Pégylation	x 1,4-1,7	Commercialisé (Etats-Unis)
	Eloctat (rFVIII-Fc)	Protéine fusion		Commercialisé
	rFVIII-SingleChain	FVIII chaîne unique		Phase II/III
Hémophilie B	N9-GP	Glycopégylation		Phase III
	Alprolix	Protéine fusion	x 2-5	Commercialisé (Etats-Unis)
	Idelvion	Protéine fusion		Commercialisé

**Tableau 2 : Liste des molécules à demi-vie prolongée pour le traitement des hémophilies A et B**

Les principaux effets indésirables liés aux traitements substitutifs sont les contaminations virales (concentrés de facteurs d'origine plasmatique uniquement) et le développement d'inhibiteurs. Le risque de contamination virale est actuellement très bien contrôlé. En effet, avec la mise en place de méthodes de purification (cryoprécipitation, chromatographie d'échange d'ions) et d'inactivation virale (traitement par solvant/détergent et nanofiltration) lors de la production de concentrés de facteurs VIII ou IX, aucun cas de contamination virale n'a été recensé ces 30 dernières années (10).

La complication majeure du traitement substitutif est l'apparition d'inhibiteurs anti-FVIII ou anti-FIX. Les inhibiteurs apparaissent chez 25 à 30 % des patients hémophiles A sévères et chez 3 à 5% des patients hémophiles B sévères (1,11). De nombreux facteurs de risque de développement d'inhibiteurs ont été identifiés chez les hémophiles A : l'origine ethnique, les antécédents familiaux d'inhibiteurs, le type de mutation du gène *F8*, certains allèles HLA (Human Leukocyte Antigen) et les modalités de traitement (12). Les facteurs de risque chez les hémophiles B sont moins bien connus mais le type de mutation affectant le gène *F9* est un déterminant majeur du développement d'inhibiteurs (13–15). L'une des particularités cliniques des patients hémophiles B avec inhibiteurs est le risque de choc anaphylactique lors de l'administration de concentré de FIX (16). Le traitement des épisodes hémorragiques chez les patients hémophiles avec inhibiteurs dépend du taux d'anticorps. En présence d'un taux faible d'anticorps (<5 UB.ml<sup>-1</sup>) il est possible de saturer l'inhibiteur en administrant des doses importantes de concentré de facteur. En cas de concentration importante d'inhibiteurs, les agents by-passants tels que le FVIIa recombinant et les concentrés de facteurs du complexe prothrombique activés sont utilisés. Ces deux produits sont équivalents en terme d'efficacité

et d'effets indésirables thrombotiques mais la réponse à l'un ou à l'autre de ces traitements est variable en fonction des patients (17,18). Enfin, l'induction d'une tolérance immune est actuellement la meilleure méthode pour éradiquer un inhibiteur. Le taux de réussite est plus important chez les hémophiles A (succès dans 60-80% des cas) que chez les hémophiles B (succès dans moins de 50% des cas) et dépend de plusieurs facteurs dont le titre initial de l'inhibiteur (1).

La desmopressine est un traitement alternatif utilisable chez les patients hémophiles A mineurs. Cet analogue de l'hormone antidiurétique entraîne la libération des stocks endothéliaux de VWF et de FVIII (augmentation du taux de FVIII : x 3-6 en une heure). Cependant, la réponse à la desmopressine est variable en fonction des patients et une épreuve thérapeutique est nécessaire avant toute utilisation (19).

De nouvelles stratégies thérapeutiques pour le traitement de l'hémophilie sont actuellement en développement. L'une de ces stratégies est l'utilisation de molécules FVIII-mimétiques pour traiter l'hémophilie A. Un anticorps bi-spécifique (ACE910, emicizumab) reconnaissant à la fois le FIXa et le FX est actuellement à l'essai (phase III). Cette thérapeutique a pour avantage d'avoir une demi-vie longue (environ 30 jours) et d'être administrable par voie sous-cutanée (20,21). Une autre stratégie consiste à restaurer l'hémostase en inhibant les systèmes inhibiteurs soit par l'administration d'un anticorps anti-TFPI (concizumab, phase I/II) (22) soit par l'utilisation d'ARN interférants (siRNA) dirigés contre l'antithrombine (fitusiran, ALN-AT3, phase I/II) (23). Enfin, plusieurs essais de thérapie génique pour le traitement des hémophilies A et B sévères sont actuellement en cours d'évaluation. Cette approche thérapeutique a pour objectif de restaurer chez les patients une production significative de FVIII ou de FIX afin de limiter le risque de survenue d'hémorragie spontanée (24,25).

## **2. Les déficits rares de la coagulation**

Les déficits rares de la coagulation sont un ensemble de pathologies hémorragiques comprenant les déficits en FII (OMIM 613679), FV (OMIM 227400), FVII (OMIM 227500), FX (OMIM 227600), FXI (OMIM 612416), FXIII (OMIM 613235 et 613225), le déficit en facteurs vitamine K-dépendants (VKCFDs, OMIM 277450) et le déficit combiné en FV et FVIII (OMIM 227300). Leur transmission

se fait sur le mode autosomique récessif dans la très grande majorité des cas (26). Ces déficits sont rapportés dans la population générale avec des fréquences allant de 1/500 000 pour le déficit en FVII à 1/ 2 000 000 pour les déficits en FII et FXIII (26,27) (Tableau 3).

Déficit	Prévalence estimée	Gènes impliqués	Cas répertoriés dans le registre FranceCoag, nombre de patients
Prothrombine	1 / 2 000 000	<i>F2</i>	1
FV	1 / 1 000 000	<i>F5</i>	51
FVII	1/500 000	<i>F7</i>	167
FX	1 / 1 000 000	<i>F10</i>	25
FXI	1 / 1 000 000	<i>F11</i>	178
FXIII	1 / 2 000 000	<i>F13A1</i> et <i>F13B</i>	27
Déficit combiné FV et FVIII	1 / 1 000 000	<i>LMAN1</i> et <i>MCFD2</i>	13
Déficit en facteurs vitamine K-dépendants	<50 familles rapportées	<i>GGCX</i> et <i>VKORC1</i>	Aucun

**Tableau 3 : caractéristiques générales des déficits rares de la coagulation (28).** (<https://www.francecoag.org>, Octobre 2017).

### 1. Présentation clinique

La présentation clinique est très variable en fonction du type de déficit mais aussi entre les individus porteurs d'un même déficit. Les formes cliniques vont de l'apparition de saignements mineurs post-traumatiques à des saignements sévères dès la naissance. Les principaux signes cliniques associés aux différents déficits sont répertoriés dans le Tableau 4. Les individus hétérozygotes sont asymptomatiques le plus souvent. Des complications gynécologiques telles que des ménorragies, des avortements spontanés et des hémorragies de la délivrance sont observées chez environ 20 % des femmes porteuses d'un de ces déficits (29,30).

Dans les déficits rares de la coagulation, le taux résiduel de facteur n'est pas toujours corrélé au risque hémorragique. La symptomatologie clinique est assez bien corrélée au taux résiduel de facteur pour les déficits en FV+FVIII, FX et FXIII. Cette corrélation est très faible voire inexistante dans le cas des déficits en FV, FVII et FXI (31).

Type d'hémorragies	Hématome hémarthrose	Post-traumatique	Digestive	Cavité buccale	Ménorrhagie	Epistaxis	Chute du cordon	SNC
Déficits	FII / FX	FII / FV / FVII / FV+FVIII / FX / FXI / FXIII / VKCFD	FII / FV	FV+FVIII / FXI / FXIII	FII / FV / FV+FVIII / FVII / FX / FXI / FXIII	FV / FV+FVIII / FVII / FX	FX / FXIII / VKCFD	FXIII / VKCFD

**Tableau 4 : Signes cliniques les plus fréquents associés aux déficits rares de la coagulation (28).** SNC : système nerveux central.

## 2. Traitement des déficits rares de la coagulation

La prise en charge des déficits rares de la coagulation repose essentiellement sur l'administration d'un traitement substitutif lors d'un épisode hémorragique. Des concentrés de facteur spécifiques ne sont disponibles que pour les déficits en FVII, FXI et XIII. Le traitement des autres déficits est basé soit sur l'administration de concentrés du complexe prothrombique (déficit en FII, FX et VKCFD) ou de plasma frais congelé (déficit en FV, FVIII+FV). Chez certains patients présentant un phénotype sévère, un traitement prophylactique peut être envisagé (32). Le développement d'inhibiteurs suite à l'administration d'un traitement substitutif a été rapporté dans des cas de déficits en FII, FVII, FXI et FXIII (33–37).

### **Partie 3 : Génétique des déficits héréditaires en facteur de la coagulation : exemple de l'hémophilie A.**

Depuis le séquençage du gène *F8* en 1984, une très grande diversité d'anomalies génétiques affectant ce gène a été mise en évidence et plus de 2000 mutations sont actuellement répertoriées dans les bases de données internationales (FVIII variant database, <http://www.factorviii-db.org/>; CHAMP Mutation List Database, [www.cdc.gov/ncbddd/hemophilia/champs.html](http://www.cdc.gov/ncbddd/hemophilia/champs.html)). En France, la très grande majorité des hémophiles A est génotypés. L'intérêt principal de la réalisation d'une étude génétique chez les hémophiles A est la détection de la mutation responsable afin de pouvoir identifier avec certitude les femmes conductrices. Cependant dans certains cas, les données de génétique peuvent fournir des informations importantes pour la prise en charge des patients (risque de développer un inhibiteur, réponse au traitement par la desmopressine).

Le spectre des mutations responsables d'HA est différent en fonction de la sévérité du phénotype. Cinquante pourcents des anomalies trouvées chez les hémophiles A sévères sont des grands réarrangements de type inversion de l'intron 22 (Inv22) ou de l'intron 1(Inv1). Des délétions ou insertions de taille variable sont présentes dans 20 % des cas et les mutations ponctuelles représentent environ 25% des cas (15% de faux-sens et 10 % de non-sens). A l'inverse, les mutations faux-sens représentent 75% des anomalies responsables d'HA modérée et mineure (38).

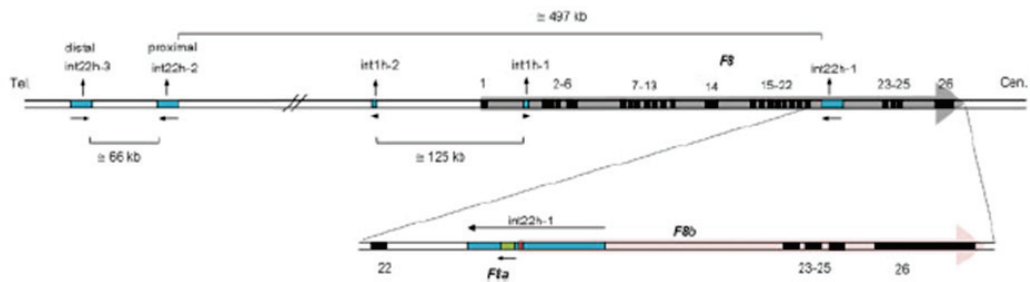
#### **1. Structure du gène *F8***

Le gène *F8* est localisé au niveau du bras long du chromosome X (Xq28, chrX:154,064,064-154,250,998, hg19, <https://genome.ucsc.edu>). C'est un gène de grande taille s'étendant sur 189 kb, composé de 26 exons de taille variable allant de 69 pb (exon 5) à 3016 pb (exon 14). La taille des introns varie aussi beaucoup et va de 200 pb (intron 17) à 32 kb (intron 22). Au total 6 introns ont une taille supérieure à 14 kb (introns 1, 6, 13, 14, 22 et 25 ; Tableau 5) (39,40).

Exon	Taille, pb	Intron	Taille, pb
1	313	1	22,9
2	122	2	2,6
3	123	3	3,9
4	213	4	5,4
5	69	5	2,4
6	117	6	14,2
7	222	7	2,6
8	262	8	0,3
9	172	9	4,8
10	94	10	3,8
11	215	11	2,8
12	151	12	6,3
13	210	13	16,0
14	3106	14	22,7
15	154	15	1,3
16	213	16	0,3
17	229	17	0,2
18	183	18	1,8
19	117	19	0,6
20	72	20	1,6
21	86	21	3,4
22	156	22	32,4
23	145	23	1,4
24	149	24	1,0
25	177	25	22,4
26	1958		

**Tableau 5 : Taille des exons et des introns composant le gène *F8*.**

L'intron 22 contient une région riche en CpG qui sert de promoteur bidirectionnel pour deux gènes, *F8A* et *F8B*, nichés dans le gène *F8* (41,42). Le gène *F8A*, ne contenant aucun intron, est transcrit dans le sens opposé du gène *F8* et s'étend sur 2 kb dans l'intron 22. Il est exprimé dans de nombreux tissus et code une protéine de 40 kDa (Huntingtin-associated protein 40) qui pourrait avoir un rôle dans la maladie de Huntington (43,44). Le gène *F8B* est transcrit dans la même direction que le gène *F8*, son premier exon (codant 8 acides aminés) est localisé dans l'intron 22 du gène *F8* et les autres exons correspondent aux exons 23 à 26 du gène *F8*. La fonction du gène *F8B* est encore mal connue (45). Le promoteur de ces deux gènes est localisé dans une région de 9,5kb appelées région homologue de l'intron 22-1 (*int22h-1*). Des copies homologues de cette région sont localisées approximativement 300 et 400 kb en amont du gène *F8* (*int22h-2* et *int22h-3* ; Figure 2). Ces régions homologues sont impliquées dans des recombinaisons intrachromosomiques appelées Inv22.



**Figure 2 : Structure du gène *F8* (46).** Les flèches noires indiquent le sens de la transcription du gène *F8A* ainsi que les orientations des régions homologues de l'intron 22 (*int22h*) et de l'intron 1 (*int1h*). Les flèches grise et rose indiquent le sens de transcription des gènes *F8* et *F8B*. Les exons du gène *F8* sont en noirs. Les copies intragéniques et extragéniques des régions homologues des introns 22 et 1 sont en bleu.

L'intron 1 du gène *F8* contient une région d'environ 1kb appelée *int1h-1*. Cette région possède une copie homologue (*int1h-2*) localisée 125 kb en amont du gène *F8*. Ces deux séquences homologues sont impliquées dans des recombinaisons intrachromosomiques appelées Inv1.

Le gène *F8* est transcrit en un ARNm de 9,01 kb codant une protéine de 2351 acides aminés composée d'un peptide signal et d'une succession de domaines protéiques fonctionnels (A1-a1-A2-a2-B-a3-A3-C1-C2 ; Tableau 6 et Figure 3).

Domaine	Positions, acides aminés
Peptide signal	1-20
A1	20-355
a1	356-390
A2	391-729
a2	730-759
B	760-1667
a3	1668-1708
A3	1709-2038
C1	2039-2191
C2	2192-2351

**Tableau 6 : Liste des différents domaines protéiques fonctionnels composant le FVIII.**



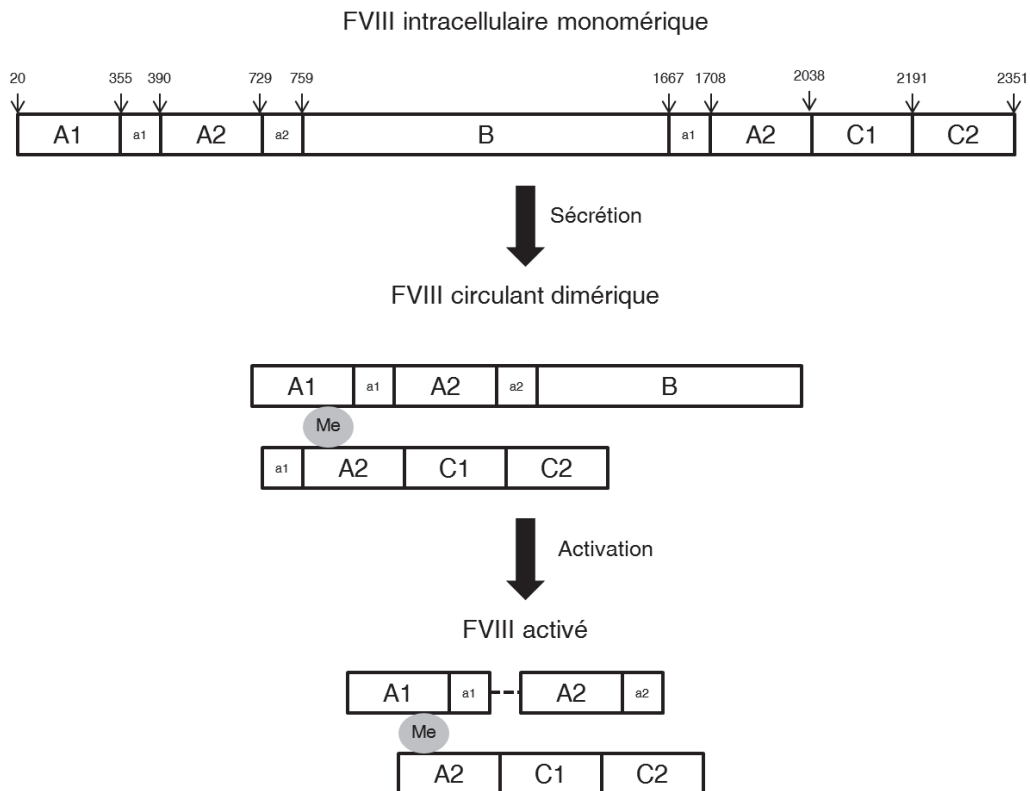


Figure 3: Structure du FVIII intracellulaire, circulant et du FVIII activé.

## 2. Origine des mutations du gène *F8*

Lors de la naissance d'un hémophile A sévère, l'absence d'antécédents familiaux d'hémophilie est observée dans 50 % des cas. Les cas sporadiques représentent environ 30% des phénotypes modérés et mineurs. Ces cas sporadiques sont liés soit à l'apparition d'une mutation *de novo* chez l'enfant soit à l'ignorance de la mère de son statut de conductrice (47). Les études familiales de ces cas sporadiques ont montré que dans presque 90% des cas, la mutation est apparue une ou deux générations avant (88% des mutations retrouvées chez la mère et seulement 18% chez la grand-mère) (47). L'importance des cas sporadiques d'HA est liée à l'existence d'une probabilité élevée de mutation du gène *F8*. Cette probabilité est estimée à  $4,2 \cdot 10^{-5}$  par génération de cellules germinales ce qui est par exemple 10 fois supérieure à la probabilité de mutation du gène *F9* (48). Ce taux de mutation permet d'expliquer que l'incidence de l'HA est à la fois la même dans le monde entier et indépendante de l'accès des hémophiles aux traitements et au diagnostic prénatal (46).

Les phénomènes d'Inv22 et d'Inv1, qui représentent 50% des cas d'HA sévère, ont lieu essentiellement lors de la méiose des gamètes mâles. En effet, du fait de sa petite taille, le chromosome Y ne peut stabiliser le chromosome X ce qui facilite ces phénomènes de recombinaison intrachromosomiques (49). Les mutations ponctuelles sont aussi prédominantes chez les gamètes mâles du fait de la persistance de la spermiogénèse tout au long de la vie. A l'inverse, les grandes délétions, conséquences de recombinaisons interchromosomiques, ont lieu essentiellement dans les gamètes femelles (50).

Une autre observation intéressante dans l'HA est la proportion importante de mosaïque somatique (liée à l'apparition de mutation au début de l'embryogénèse). Environ 25% des mutations ponctuelles sont associées à un phénomène de mosaïque somatique (51). Ce phénomène semble plus rare pour les inversions, en effet, un seul cas de mosaïque somatique associée à une inv22 a été rapporté dans la littérature (52). Ce phénomène a une grande importance lors de la détermination du statut de conductrice des femmes. Une mosaïque germinale peut masquer le fait que la femme est conductrice d'HA.

### **3. Les mutations du gène *F8***

#### **1. Les réarrangements génomiques**

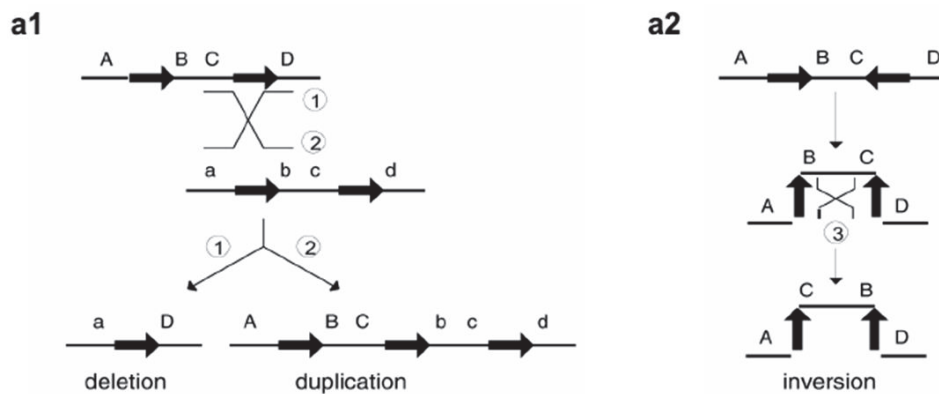
Sous le terme réarrangements génomiques sont classées les délétions, duplications et inversions de grande taille (de quelques centaines à plusieurs millions de nucléotides) (53). Les réarrangements génomiques peuvent être récurrents (même taille, mêmes points de cassure) ou non (taille différente, points de cassure différents).

#### **a) Réarrangements récurrents**

#### **La recombinaison homologue non allélique**

La majorité des réarrangements génomiques récurrents sont causés par des recombinaisons homologues non alléliques (NAHR pour Non allelic Homologous Recombination) entre deux séquences répétées du génome appelées LCRs (low-copy repeats). Les LCRs sont des régions de 10 à 300 kb présentant plus de 95% d'homologie entre elles (54). Du fait de leur grande homologie de

séquence, les LCRs peuvent s'aligner lors des phénomènes de méiose ou de mitose. Ces erreurs d'alignements peuvent entraîner des réarrangements génomiques dans les cellules filles suite à un crossing-over. Quand les deux LCRs localisées sur le même chromosome sont orientées dans le même sens, la NAHR entre ces deux régions est responsable de délétion et/ou duplication. En revanche s'ils sont orientés en sens inverse, la recombinaison de ces deux éléments est responsable d'un phénomène d'inversion (Figure 4). Les recombinaisons entre deux LCRs localisées sur des chromosomes différents sont responsables de translocations chromosomiques (55).

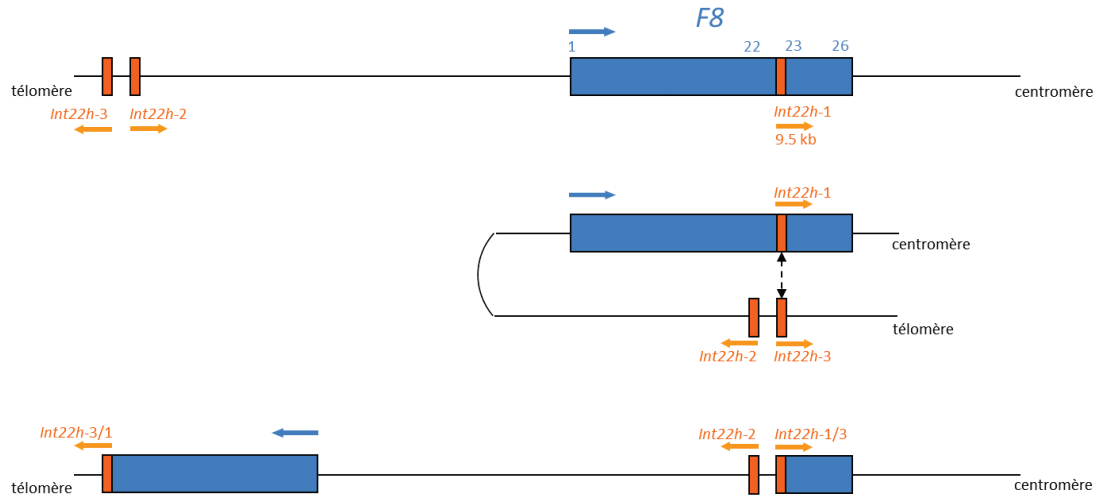


**Figure 4 : Réarrangements génomiques résultant d'une recombinaison homologue non allélique entre deux séquences répétées LCRs (53).** Les LCRs sont représentées sous forme de flèches noires dont le sens indique l'orientation du LCR. Les lettres représentent des régions de l'ADN flanquant les LCRs. a1 : recombinaison entre deux LCRs orientées dans le même sens. a2 : recombinaison entre deux LCRs orientées en sens inverse.

### Les réarrangements récurrents dans l'hémophilie A

La moitié des cas d'HA sévères sont liés à un réarrangement de type NAHR. En effet, Les Inv22 et Inv1 sont responsables respectivement de 43% et 2,3% des HA sévères (56–59). L'Inv22 est le résultat d'une inversion entre une LCRs (*int22h-1*) localisée dans l'intron 22 du gène *F8* et l'une des deux copies extragéniques de cette LCRs (*int22h-2* et *int22h-3*) situées respectivement 300 et 400 kb en amont du gène *F8* (56,57,60). *Int22h-2* a la même orientation qu'*int22h-1* alors qu'*int22h-3* est dans le sens opposé. Il y a deux type d'Inv22 : le type 1 (Inv22-1) impliquant directement *int22h-1* et *int22h-3* et le type 2 (Inv22-2) impliquant *int22h-1* et *int22h-2* après une inversion préliminaire entre *int22h-2* et *int22h-3* (61). La faible fréquence de l'inversion préliminaire entre *int22h-2* et *int22h-3* peut expliquer que l'Inv22-2 ne représente que 16% des cas et l'Inv22-1 en représente 80% (62,63). Enfin, les 4% restant sont des Inv22 plus complexes associant la duplication de la région contenant

les séquences *int22h-2* et *int22h-3* et une inversion (64). Le phénomène d'Inv22 est responsable d'une cassure du gène *F8* avec la séparation des exons 1-22 et des exons 23-26 (Figure 5). Depuis la découverte de ce phénomène en 1993, différentes méthodes de diagnostic ont été développées telles que le Southern Blot (SB) et les PCR long-fragment (LR-PCR) (56,63).



**Figure 5 : L'inversion de l'intron 22 de type 1 (65).** Le phénomène d'inversion de l'intron 22 est la conséquence d'une recombinaison homologue non allélique intrachromosomique entre deux séquences homologues (*int22h-1* et *int22h-3*) orientées en sens inverse.

L'Inv1, initialement décrite par Brinke *et al.* (66) puis par Bagnall *et al.* (67), est le résultat d'un mécanisme similaire à l'Inv22 impliquant une séquence de 1041 pb (*int1h-1*) localisée dans l'intron 1 du gène *F8* et une copie extragénique de cette séquence (*int1h-2*) localisée approximativement à 140kb en amont du gène *F8* et orientée en sens inverse. Ce phénomène est aussi responsable d'une cassure dans le gène *F8* avec la séparation de l'exon 1 et des exons 2-26. La recherche d'une Inv1 chez les patients est réalisée par PCR (67).

Un troisième type d'inversion entre une séquence localisée dans l'intron 1 différente de *int1h-1* et une copie homologue extragénique a été décrite dans un cas d'HA sévère (68).

Quelques cas de duplications et de délétions de la région localisée entre les séquences *int22h-1* et *int22h-2* (0,5 Mb) ont aussi été décrits dans la littérature (69,70). Ces phénomènes sont la conséquence de NAHR entre les LCRs *int22h-1* et *int22h-2* localisées sur deux chromatides différentes. La duplication de cette région de 0,5Mb a été observée chez des hommes non hémophile. La délétion

réciroque n'a, à ce jour, été décrite que chez des femmes et aurait une fréquence allélique de 0,21% dans la population générale (69,71).

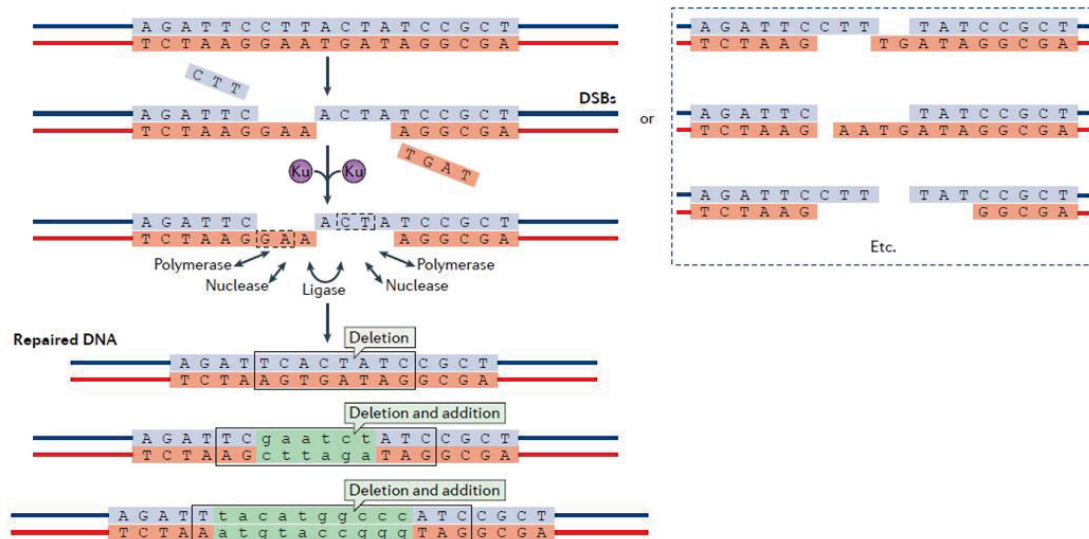
b) Réarrangements non récurrents

A la différence des réarrangements récurrents, les anomalies non récurrentes possèdent des points de cassure uniques et ont des tailles variables. Les points de cassures des réarrangements non récurrents sont caractérisés soit par l'absence totale d'homologie soit par la présence de microhomologies de séquences (4 à 20 nucléotides). Parfois, des insertions ou délétions de quelques paires de bases sont présentes au niveau de ces points de cassure. Les mécanismes responsables de ces anomalies peuvent être classés en deux catégories : ceux ne nécessitant pas d'étape de réplication de l'ADN et ceux nécessitant une étape de réplication (72).

**Mécanismes sans réplication**

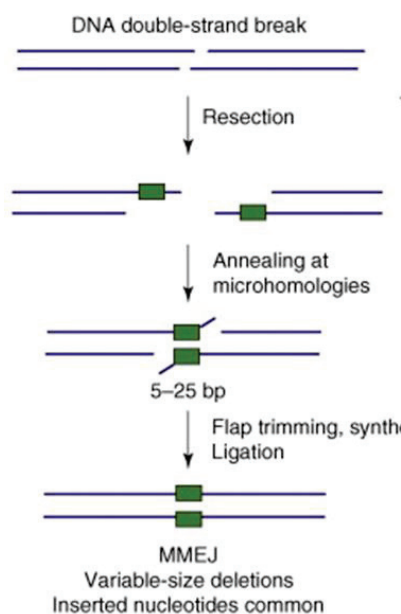
Ce type de réarrangements est souvent la conséquence de l'action de mécanismes de réparation de l'ADN suite à une cassure double brin. Deux mécanismes principaux de réparation de l'ADN après une cassure double brin ont été décrits : la jonction d'extrémités non-homologues (NHEJ pour Non Homologous End-joining) et la jonction d'extrémités présentant des microhomologies (MMEJ pour Microhomology-mediated End-joining).

La NHEJ permet de lier les deux brins d'ADN après une cassure en absence d'homologie de séquence (Figure 6). La ligation des deux brins d'ADN peut entraîner, au niveau du point de cassure, la délétion de quelques nucléotides (1 à 4 le plus souvent) ou bien l'insertion de molécules d'ADN libres (tels que des rétrotransposons) (73).



**Figure 6 : Jonction d'extrémités non-homologues NHEJ (74).** Suite à la cassure double brin de l'ADN, le complexe enzymatique Ku se lie à l'ADN. Ce complexe possède des activités polymérase, ligase et nucléase. Ces activités peuvent intervenir dans n'importe quel ordre afin d'éliminer ou d'ajouter des nucléotides. Plusieurs cycles d'élimination/addition de nucléotide peuvent avoir lieu. Ces étapes ont pour objectif la création de microhomologie à l'extrémité des brins d'ADN afin de faciliter l'étape de ligation. Ce mécanisme est souvent responsable de l'apparition d'erreur au niveau du point de cassure mais peut aussi lier les deux brins d'ADN sans modification de la séquence.

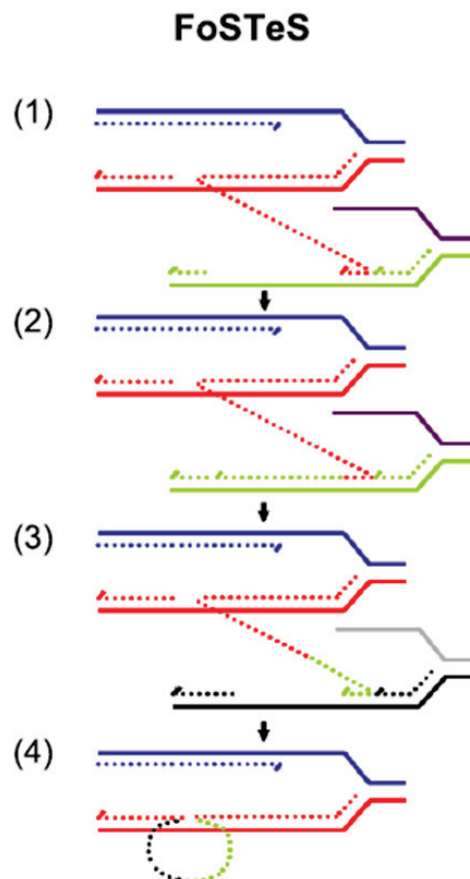
Lors de la MMEJ, avant ligation, les deux brins d'ADN s'hybrident au niveau de séquences homologues de 5 à 25 bp situées dans l'entourage du point de cassure (Figure 7). Ce mécanisme de réparation est fréquemment responsable de la délétion de la région d'ADN localisée entre les deux séquences homologues (72).



**Figure 7 : Jonction d'extrémités présentant des microhomologies (MMEJ) (75).** Après une cassure double brin de l'ADN, des nucléotides sont éliminés afin de permettre l'alignement de deux courtes séquences homologues de 5 à 25 pb.

## Mécanismes avec réplication: Microhomology-Mediated Break-Induced Replication (MMBIR) / Fork Stalling and Template Switching (FoSTeS)

Ces deux mécanismes sont très proches et ont lieu lors d'arrêt de la fourche de réplication (soit du fait d'une cassure simple brin de l'ADN en réplication soit lié à la structure secondaire de l'ADN). Le brin d'ADN libre s'hybride au niveau d'une séquence présentant des microhomologies localisées à un autre endroit sur la molécule d'ADN ce qui permet la création d'une nouvelle fourche de réplication. Ce phénomène peut se répéter à plusieurs reprises et être à l'origine de réarrangements très complexes associant duplication, tripllication, délétion et/ou inversion (Figure 8). Ce type de réarrangements est caractérisé par la présence au niveau de chaque point de cassure de microhomologies de séquence.



**Figure 8 : Fork Stalling and Template Switching (FoSTeS).** Suite à l'arrêt de la fourche de réplication. Le brin d'ADN libre va s'hybrider sur une séquence présentant une microhomologie (1). Cette hybridation permet la création d'une nouvelle fourche de réplication (2). Le même mécanisme peut se répéter plusieurs fois (3). Finalement, la réplication reprend sur le brin initial (4).

## Les réarrangements non récurrents dans l'hémophilie A

Les délétions de grande taille localisées dans le gène *F8* sont responsables de 3% des HA sévères (38) et plus de 255 de ces délétions sont répertoriées dans les bases de données internationales. Dans quelques cas, les points de cassure ont été caractérisés. La majorité de ces délétions sont causées par des mécanismes de type NHEJ et MMBIR (76–81). Un réarrangement complexe du gène *F8* associant à la fois une Inv1 et une délétion (2,56 kb) / insertion (256 kb) lié à un mécanisme de type MMBIR / FoSTeS a aussi été rapporté dans la littérature (82).

De nombreuses duplications (comprenant un à plusieurs exons du gène *F8*) ont été rapportées dans la littérature. Elles ont été identifiées majoritairement chez des hémophiles A sévères mais aussi chez des hémophiles modérés et mineurs. Les points de cassure ont été caractérisés pour plusieurs d'entre elles et des microhomologies ou des séquences répétées (séquences *Alu* par exemple) ont été fréquemment mises en évidence (83). Ces duplications pourraient être la conséquence de réarrangements de type MMEJ ou MMBIR / FoSTeS. Un cas particulier est la duplication de l'exon 13 du gène *F8* qui est trouvée chez environ un tiers des hémophiles A mineurs originaires du Nord de l'Italie (84). La fréquence importante de cette anomalie dans cette région est liée à un effet fondateur. Cette duplication est le résultat d'une recombinaison non homologe dont le mécanisme exact est non déterminé à ce jour (85). La présence d'un épissage alternatif chez ces patients (saut de l'exon 13 surnuméraire) permet la production d'une faible quantité de FVIII normal (86).

### 2. Les délétions et insertions de petite taille

Des petites délétions et insertions (< 50 pb) sont retrouvées dans 16% des HA sévères (38). La majorité de ces anomalies sont la conséquence de MMEJ et NHEJ ayant eu lieu suite à une cassure double brin de la molécule d'ADN. Ces insertions entraînent souvent un décalage du cadre de lecture et l'apparition d'un codon stop prématuré. Une exception est la mutation c.6658\_6660del (p.Ala2220del) qui est responsable d'HA mineure à modérée (87). Cette délétion conserve la cadre de lecture mais entraîne une diminution de l'affinité du FVIII pour le VWF et les phospholipides.



Les petites délétions ou insertions (1-2 pb) peuvent apparaître suite à des erreurs de réplication au niveau de régions composées d'une répétition d'un seul nucléotide. L'exon 14 du gène *F8* contient plusieurs répétitions de A particulièrement sensibles à ce phénomène (c.3629\_3637 et c.4372\_4379) (88–90). Bien que ces mutations entraînent un décalage du cadre de lecture, certaines sont responsables d'un phénotype modéré d'HA. Ce phénomène est expliqué par l'existence d'erreurs de transcription au niveau des mêmes régions répétées lors de la synthèse de l'ARNm permettant la restauration du cadre de lecture (91,92) et la production d'une petite quantité de FVIII actif.

### 3. Les substitutions nucléotidiques

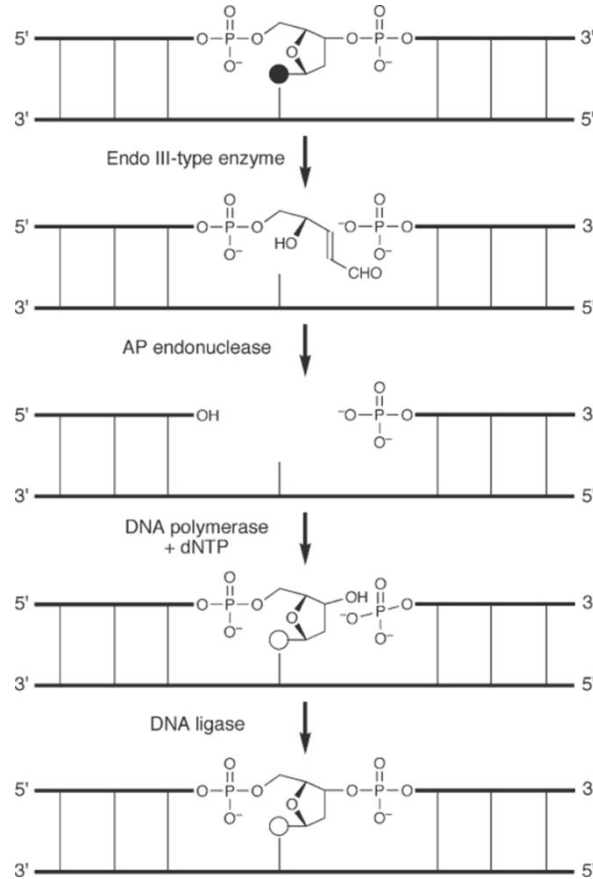
#### a) Origine des substitutions nucléotidiques

Lors de la réplication de l'ADN, l'ADN polymérase a un taux d'erreur estimé à  $10^{-4}$ - $10^{-5}$  par nucléotide. Ce taux d'erreur est diminué à  $10^{-7}$  par nucléotide du fait de l'activité exonucléase de certaines ADN polymérases. Il existe un système de réparation des erreurs de l'ADN polymérase chez l'Homme (MMR pour DNA Mismatch Repair). Ce système permet de diminuer le nombre d'erreur de réplication à  $10^{-10}$ /nucléotide (93). Le génome humain contenant  $6.10^9$  nucléotide, en moyenne il y a donc moins d'une erreur de réplication par division cellulaire chez l'Homme. Les protéines formant le système MMR reconnaissent les mésappariements sur l'ADN double brin, possèdent une activité exonucléase pour éliminer les nucléotides erronés, une activité polymérase pour synthétiser un nouveau brin d'ADN et une activité ligase (93).

Des agressions chimiques et/ou physiques peuvent aussi modifier les bases de l'ADN et être responsables de mutation. Il existe deux principaux mécanismes de réparation de ce type de lésion : le système de réparation par excision de bases (BER) et le système de réparation par excision de nucléotides (NER).

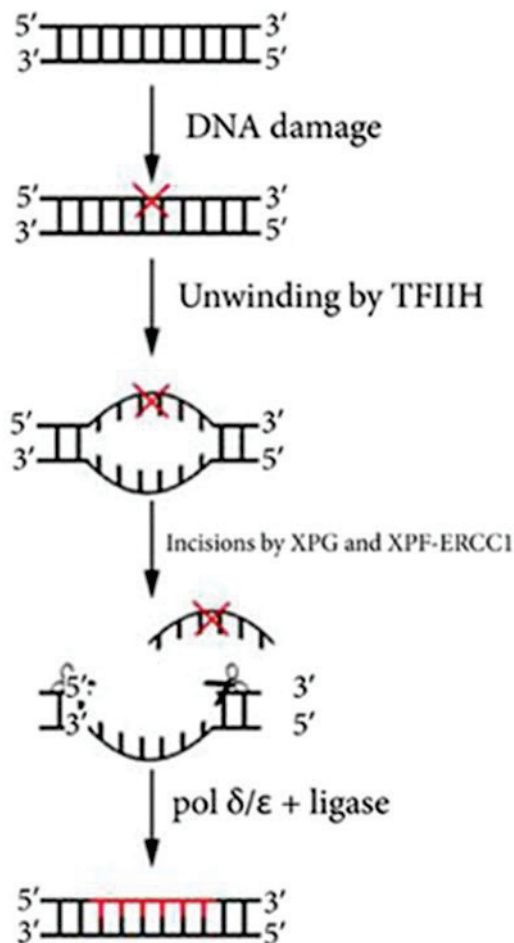
Le système BER permet d'éliminer et de remplacer les bases modifiées suite à une agression chimique (alkylation, oxydation, désamination). Dans une première étape, une DNA glycosylase spécifique reconnaît la base modifiée (il existe différentes glycosylases spécifiques pour chaque modification ex : 8-oxoguanine DNA glycosylase, Uracile N-glycosylase) et coupe la liaison entre la

base et le désoxyribose créant un site apurinique ou apyrimidique (AP) dans la molécule d'ADN (Figure 9). Le site AP est ensuite clivé par l'apurinique endonucléase 1 puis le nucléotide correct est inséré grâce aux actions successives d'une ADN polymérase et d'une ligase (94).



**Figure 9 : Réparation de l'ADN par excision de base (BER) (95).** Dans une première étape, une DNA glycosylase spécifique reconnaît la base modifiée et coupe la liaison entre la base et le désoxyribose créant un site apurinique ou apyrimidique (AP) dans la molécule d'ADN. Le site AP est ensuite clivé par l'apurinique endonucléase 1 puis le nucléotide correct est inséré grâce aux actions successives d'une ADN polymérase et d'une ligase

Le système NER a pour fonction de réparer des lésions étendues de l'ADN (suite à des radiations par exemple). Ce système fonctionne en trois étapes : la reconnaissance de la région lésée, l'excision des nucléotides lésés et la synthèse du nouveau brin d'ADN. Le système NER comprend deux voies, l'une spécifique des régions non codantes de l'ADN (Global Genome NER) et l'une spécifique des régions transcrites (transcription coupled-NER) (96).



**Figure 10 : Réparation de l'ADN par excision de nucléotide (NER) (97).** Après reconnaissance de la région de l'ADN lésée, la séquence d'ADN est éliminée puis un nouveau brin est synthétisé.

Malgré ces différents systèmes permettant d'assurer l'intégrité du génome d'une cellule, certaines erreurs / lésions ne sont pas corrigées avant le nouveau cycle cellulaire. Ces erreurs ne sont plus reconnues comme telles dans les cellules filles et sont transmises aux générations suivantes. Les mutations ponctuelles correspondent au remplacement d'un nucléotide par un autre (substitution nucléotidique). Elles représentent 25 % des mutations responsables de phénotypes sévères d'HA et plus de 75% des mutations responsables des phénotypes modérés et mineurs (38).

Il est intéressant de remarquer qu'un tiers des substitutions sont localisées au niveau des dinucléotides CG (98). La fréquence anormale de substitution au niveau de ces structures est liée au phénomène de méthylation des îlots CpG de l'ADN. En effet la désamination de la méthylcytosines génère une thymine. La thymine étant une base normale de l'ADN, l'erreur n'est pas détectée par le

système BER. Ainsi, la désamination spontanée des méthylcytosine entraîne l'apparition de substitution C>T (ou G>A en fonction du brin d'ADN où est localisée la méthylcytosine). Le gène *F8* contient 70 îlots CpG et environ 35% des substitutions décrites sur ce gène sont localisées au niveau de ces îlots (99).

#### b) Mécanismes physiopathologiques des substitutions du gène *F8*

##### **Les variations de la région promotrice**

Le promoteur du gène *F8* s'étend sur 1175 pb (position c.-1175 à c.-1). Il est composé de plusieurs sites de liaison de facteur de transcription tels que HNF 1 et 4 (Hepatic Nuclear Factor 1 et 4), NF-κB (Nuclear factor κB) et C/EBPα et β (CCAAT/enhancer-binding protein). Des études d'expression ont montré que seule la région située entre les positions c.-279 à c.-64 contenant les sites de liaison de NF-κB et C/EBPα et β est indispensable à l'activité promotrice (100).

Un petit nombre de variations localisées dans la région promotrice du gène *F8* ont été décrites à ce jour (101–104). Ces anomalies sont rencontrées dans tout type de sévérité et sont associées à une mauvaise réponse des patients au traitement par la desmopressine (105,106). Des études *in vitro* utilisant un modèle de gène rapporteur ont démontré la pathogénicité de certaines de ces variations (103,104). Seules les anomalies localisées dans la région c.-279 à c.-64 du promoteur sont délétères. Les variants localisés en dehors de cette région semblent être des polymorphismes rares (c.-51G>A, c.-664G>A) (104).

##### **Les variations des régions exoniques**

Les substitutions exoniques peuvent être classées en trois catégories :

- les substitutions non-sens : elles entraînent l'apparition d'un codon stop (TAA, TAG, TGA) et sont responsables de la synthèse d'une protéine tronquée non fonctionnelle. Elles sont mises en évidence dans environ 10% des HA sévères (38).

- les substitutions synonymes : du fait de la redondance du code génétique, ce type de variation n'entraîne pas de changement de l'acide aminé sur la protéine. Cependant ces anomalies peuvent avoir un effet délétère sur l'épissage du gène *F8* et donc être responsables d'hémophilie A.

- les substitutions faux-sens : elles entraînent le remplacement d'un acide aminé par un autre sur la protéine. Elles peuvent avoir des conséquences très variables sur la molécule de FVIII. On les retrouve dans tous les types de sévérité et elles sont localisées dans les domaines A et C du FVIII (Figure 11). Elles représentent respectivement 15% et 75% des mutations trouvées chez les hémophiles A sévères et non sévères (38). La pathogénicité des variations faux-sens localisées dans la région de l'exon 14 du gène *F8* codant le domaine B est très incertaine. En effet le domaine B n'est pas nécessaire à l'activité pro-coagulante du FVIII (107). Plusieurs équipes ont étudié à l'aide de modèles d'expression cellulaire de variant l'impact des substitutions faux-sens localisées dans le domaine B sur l'activité ou la biosynthèse du FVIII (108–110). La grande majorité de ces variations semblent n'être que des polymorphismes rares.

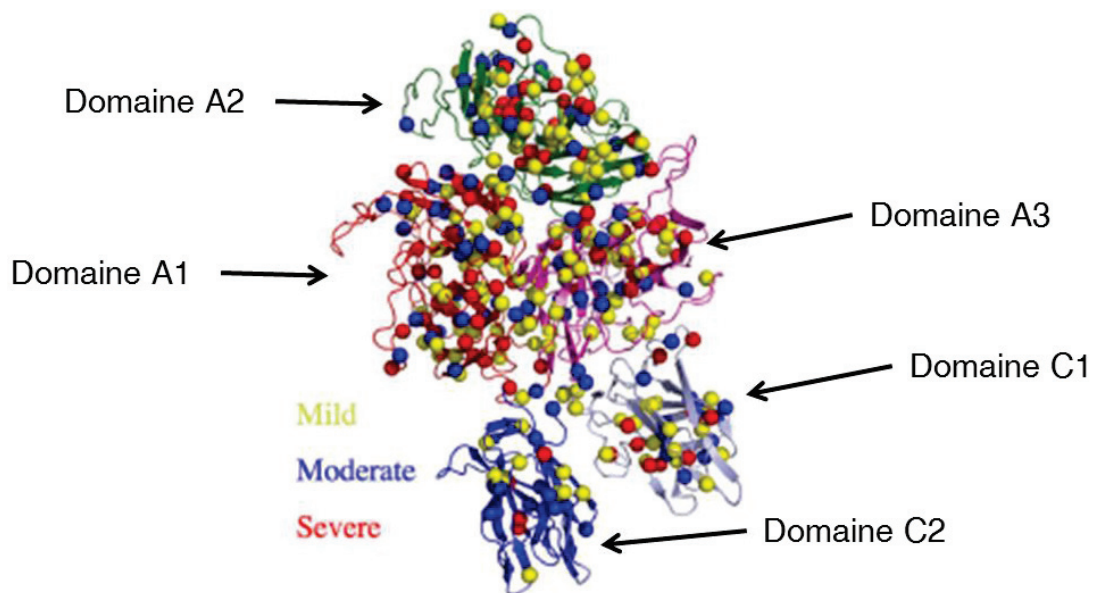


Figure 11 : Localisation des mutations faux-sens sur le facteur VIII délété en domaine B (111).

Plusieurs mécanismes physiopathologiques par lesquels les variations faux-sens sont responsables d'anomalies de biosynthèse ou d'activité du FVIII ont été décrits.

#### Altération de la structure tertiaire de la protéine de FVIII entraînant sa dégradation intracellulaire

La majorité des variations faux-sens ont un effet délétère sur la structure tridimensionnelle du FVIII. Ceci a pour conséquence une diminution importante de la sécrétion de FVIII. En effet, les protéines n'ayant pas une conformation correcte sont séquestrées dans le réticulum endoplasmique et orientées vers la voie du protéasome pour y être dégradées. La grande majorité des mutations faux-sens présentes dans les HA sévères sont responsables d'anomalies de structure du FVIII.

#### Altération de la liaison du FVIII au VWF dans la circulation sanguine.

Dans la circulation sanguine, le FVIII est lié au VWF. Ce dernier protège le FVIII d'une dégradation précoce. Certaines variations localisées dans les domaines C1 et C2 du FVIII sont responsables d'une diminution de l'affinité du FVIII pour le VWF et entraînent donc un raccourcissement de la demi-vie du FVIII (112–114). En dehors de ces deux domaines, la mutation p.Tyr1699Phe localisée dans le domaine acide a3 diminue aussi la liaison du FVIII au VWF (115). Ces mutations sont observées dans des phénotypes non sévères d'HA.

#### Altération de l'activation du FVIII par la thrombine

Trois clivages protéolytiques par la thrombine au niveau des arginines 391, 759 et 1708 sont nécessaires à la transformation du FVIII en FVIIIa. Les régions acides a1, a2 et a3, localisées en amont des sites de clivage jouent un rôle important dans l'interaction entre la thrombine et le FVIII. Des mutations localisées dans les régions acides telles que p.Tyr365Cys et p.Glu739Lys altèrent l'activation du FVIII par la thrombine (116). Ces mutations sont responsables de discordances entre les différentes techniques de dosage de l'activité plasmatique du FVIII (FVIII:C < FVIII:Cr). Ce phénomène pourrait être lié à la pré-incubation du FVIII avec la thrombine dans le test chromogénique test dit en deux temps). Il n'y a pas de phase de pré-incubation dans le test coagulométrique (test dit en un temps).

### Diminution de la stabilité du FVIII

Des variations localisées au niveau des interfaces entre les domaines A1/A2, A2/A3 et A1/A3 sont trouvées chez des hémophiles non sévères présentant une discordance importante entre les dosages chromogéniques et coagulométriques du FVIII (FVIII:C > FVIII:Cr). Ces mutations entraînent une dégradation accélérée du FVIIIa en déstabilisant les liaisons non covalentes existantes entre les différents domaines A. Il a ainsi été montré que la mutation p.Arg531His était responsable d'une dissociation spontanée du domaine A2 4 fois plus rapide par rapport au FVIII non muté (117). A ce jour, plus d'une vingtaine de mutations responsables de discordance FVIII:C >FVIII:Cr ont été décrites (118–120).

### Diminution de l'affinité du FVIIIa pour les phospholipides et le FIXa

Le FVIIIa se lie aux phospholipides chargés négativement présents à la surface des plaquettes activées via son domaine C2. Des mutations présentes dans ce domaine, comme par exemple les variants p.Ala2220Pro et p.Arg2323Cys, sont responsables d'une perte d'affinité du FVIIIa pour les phospholipides. Ces mutations sont responsables de phénotypes non sévères d'HA (121).

De la même manière, certaines mutations sont responsables d'une perte d'affinité du FVIIIa pour le FIXa. C'est le cas des mutations p.Ile585Thr et p.Ser677Phe localisées dans le domaine A2. L'encombrement stérique lié à ces mutations, soit du fait de la création d'un nouveau site de N-glycosylation (p.Ile585Thr) soit du fait de l'introduction d'une chaîne latérale encombrante (p.Ser677Phe), déstabilise le complexe FIXa/FVIIIa (122).

### **Les variations des jonctions intron/exon et des régions introniques**

Les substitutions localisées au niveau des jonctions intron/exon peuvent être responsables d'anomalies d'épissage du gène *F8*. Ce type d'anomalies est observé dans tous les grades de sévérité. Les anomalies localisées au niveau des sites d'épissage du gène *F8* ont fait l'objet d'une étude approfondie dans la section 3 de ce manuscrit.

Des variations introniques profondes peuvent aussi être responsables d'anomalies d'épissage en entraînant soit la création d'un site *de novo* d'épissage soit l'activation d'un site cryptique pré-existant (58,123–125).

### **Les variations localisées dans la région 3' non transcrite du gène *F8***

Une seule variation localisée dans la région 3' non transcrite (3'UTR) du gène *F8* a été décrite à ce jour chez un hémophile A mineur (126). La mutation c.\*59G>T est responsable de la création d'un nouveau site donneur d'épissage entraînant la perte de 159 bp dans la région 3'UTR. Les auteurs ont démontré à l'aide dans un modèle *in vitro* que ce variant était responsable d'une diminution de l'expression d'un gène rapporteur et donc d'une diminution de la transcription de l'ARNm.



## **Partie 4: Maladies hémorragiques héréditaires liées à des mutations gain de fonction.**

Des syndromes hémorragiques héréditaires familiaux ont été rapportés dans la littérature chez des individus dont aucun facteur de la coagulation n'était déficitaire. Ces patients présentent des phénotypes cliniques et biologiques variables mais ont en commun une diminution importante de la génération de thrombine mesurée par le test de génération de thrombine (TGT). Des analyses génétiques réalisées chez ces patients ont permis d'identifier des mutations soit dans le gène *F5* (codant le FV) soit dans le gène *THBD* (codant la thrombomoduline). En conférant des propriétés supplémentaires aux protéines mutées, ces mutations sont responsables des phénotypes hémorragiques observés chez les patients.

### **1. Mutations gain de fonction localisées dans le gène *F5* : FV Texas et FV Amsterdam.**

En 2001, Kuang *et al* ont décrit une grande famille originaire du Texas chez laquelle plusieurs membres présentaient un phénotype hémorragique bruyant avec présence de ménorragies importantes, d'épistaxis et d'hémorragies post-traumatiques ou post-chirurgicales (127). Peu de données existent sur la prise en charge de ces accidents hémorragiques, mais il semblerait que la plupart aient été traités par administration de plasma frais congelé. Tous les patients sont porteurs de la mutation du gène *F5* c.2350A>G (p.Ser784Gly) appelée FV Texas. Cette anomalie se transmet dans la famille sur le mode autosomique dominant.

Les patients porteurs de cette anomalie présentent un allongement modéré du temps de céphaline avec activateur (TCA) et du temps de Quick (TQ). Cependant, les taux de tous les facteurs de coagulation et du fibrinogène sont normaux. Le test de génération de thrombine est perturbé. Enfin, une concentration plasmatique de TFPI $\alpha$  10 fois supérieur à la normale est observée chez les patients porteurs de la mutation (128).

La mutation c.2350A>G, localisée dans l'exon 13 du gène *F5*, entraîne l'apparition d'une variation faux-sens (p.Ser784Gly) dans le domaine B du FV. Cette anomalie génétique n'a aucun impact sur l'activité pro-coagulante du FV. La mutation c.2350G>A favorise l'utilisation d'un site

donneur cryptique d'épissage en c.2351 entraînant la production majoritaire d'un transcrit du gène *F5* plus court où la région c.2351\_c.4457 est délétée. Ce transcrit code une isoforme plus courte du FV ayant une délétion de 702 acides aminés dans son domaine B (p.Ser784\_Asp1486delinsAsn). Cette isoforme du FV (appelée FV-court), existante chez les individus sains à l'état de traces, est la forme majoritaire circulante chez les patients porteurs de la mutation FV Texas. Les concentrations élevées de TFPI $\alpha$  sont liées à l'affinité très élevée du FV-court pour le TFPI $\alpha$  ce qui entraîne une augmentation importante de la demi-vie de ce dernier (128). Les taux élevés de TFPI $\alpha$  sont responsables d'une diminution de la génération de FXa et donc de la génération de thrombine. Cette faible génération de thrombine explique le phénotype hémorragique des patients.

Récemment, une autre anomalie génétique du gène *F5*, la mutation c.2588C>G (p.Ala863Gly) a été identifiée chez une famille ayant un phénotype hémorragique proche de celui observé avec le FV Texas (129). Les mêmes anomalies biologiques (allongement des TQ et TCA et diminution de la génération de thrombine) sont mises en évidence chez ces patients et une concentration élevée de TFPI $\alpha$  est retrouvée. Cette mutation est responsable de la création d'un site donneur *de novo* d'épissage dans l'exon 13 du gène *F5* entraînant la production d'un transcrit majoritaire plus court (délétion de la région c.2588\_c.4457) codant un FV tronqué (p.Ala863\_Pro1485del) appelé FV Amsterdam. Ce FV présente les mêmes propriétés que le FV Texas (129).

## **2. Mutations gain de fonction localisées dans le gène *THBD* : c.1611C>A (p.Cys537X).**

Récemment, plusieurs équipes indépendantes ont décrit une nouvelle pathologie hémorragique chez des patients porteurs à l'état hétérozygote de la mutation du gène *THBD* c.1611C>A (p.Cys537X) codant la TM (130–133). Cette pathologie se transmet sur le mode autosomique dominant. Cette mutation entraîne la synthèse d'une protéine tronquée des 3 derniers acides aminés du domaine transmembranaire et du domaine intracytoplasmique de la TM.

Les patients porteurs de cette anomalie génétique ne présentent aucune anomalie ni sur le bilan de coagulation standard ni sur les fonctions plaquettaires. Les seuls tests perturbés sont l'index de consommation de thrombine et le test de génération de thrombine. Les patients présentent tous des

taux de TM soluble (sTM) plasmatique très élevés (de 50 à 100 fois la normale). Les caractéristiques biologiques des patients porteurs de la mutation *THBD* c.1611C>A rapportés à ce jour dans la littérature sont présentées dans le Tableau 7.

	1	2	3	4	5	6
Référence	Dargaud <i>et al.</i> (131)	Langdown <i>et al.</i> (130)	Langdown <i>et al.</i> (130)	Langdown <i>et al.</i> (130)	Burley <i>et al.</i> (133)	Burley <i>et al.</i> (133)
Sexe	Féminin	Masculin	Féminin	Masculin	Masculin	Masculin
Age, années	37	31	Non disponible	Non disponible	59	34
ETP, nM.min	61*	54*	108*	446*	316	316
ICP, %	74-91	79	68	60	Non disponible	Non disponible
sTM, ng.ml <sup>-1</sup>	1120	845	553	433	598	633
Anomalie génétique associée	Aucune	<i>THBD</i> C.1418C>T (p.Ala473Val)	Aucune	FV Leiden	Aucune	Aucune

**Tableau 7 : caractéristiques biologiques des patients porteurs de la mutation *THBD* c.1611C>A rapportés dans la littérature.** Les tests de génération de thrombine ont été réalisés sur du plasma pauvre en plaquettes en présence de 1pM de facteur tissulaire. La génération de thrombine a été évaluée chez les patients à l'aide du potentiel endogène de thrombine (ETP). \*test de génération de thrombine réalisé en présence d'un inhibiteur du système contact (CTI ; corn trypsin inhibitor). Les ETP obtenus chez des patients sains sont respectivement : 835 nM.min, 609-1242 nM.min et 1584+/-56 nM.min dans les études de Dargaud *et al.*, Langdown *et al.* et Burley *et al.* ICP : index de consommation de prothrombine, valeurs usuelles <10%. sTM : thrombomoduline soluble.

Le phénotype clinique associé à cette anomalie génétique est caractérisé par des hémorragies prolongées post-traumatiques (suite à des chirurgies ou des traumatismes minimes). Ces hémorragies sont résistantes à l'administration de plasma frais congelé, de concentrés de FIX, d'acide tranxénamique ou de desmopressine (131,133). A l'inverse, la transfusion de concentrés plaquettaires (associés ou non à des concentrés de FVIII) et l'administration de FVIIa recombinant ou de concentrés de facteurs du complexe prothrombique activés sont des thérapies efficaces pour contrôler les saignements chez ces patients (130,131,133,134).

Les mécanismes physiopathologiques responsables des taux élevés de sTM observés chez les patients sont inconnus à ce jour et ont fait l'objet d'un travail présenté dans la section 4 de ce document.

**Section 2. Anomalies génétiques de grande taille dans l'hémophilie: mise en évidence et intérêt pour le conseil génétique.**

**Partie 1. Caractérisation par analyse chromosomique sur puce à ADN de 6 délétions totales du gène *F9*: implication de la délétion du gène *SOX3* dans les déficits mentaux observés chez certains patients.**

**Study of six patients with complete F9 deletion characterized by cytogenetic microarray: role of the SOX3 gene in intellectual disability.**

J Thromb Haemost. 2016, 14: 1988-1993.

**Y. Jourdy** \*† N. Chatron ‡ M. L. Carage \*M. Fretigny \* S. Meunier ¶C. Zawadzki \*\*V. Gay ††, C. Negrier \*†¶D. Sanlaville ‡§C. Vinciguerra \*†

\* Hospices Civils de Lyon, Hôpital Edouard Herriot, Service d'hématologie Biologique, Lyon, France  
† EAM 4174 Hémostase, Inflammation et Sepsis, Université Claude Bernard Lyon 1, France.

‡ Hospices Civils de Lyon, Groupe Hospitalier Est, Laboratoire de Cytogénétique Constitutionnelle, Bron, France.

§ CRNL, équipe TIGER, INSERM U1028, CNRS UMR5292, Université Claude Bernard, Lyon1, France.

¶ Hospices Civils de Lyon, Hôpital Cardiologique Louis Pradel, Unité d'Hémostase Clinique, Bron, France.

\*\* CHRU de Lille, Laboratoire d'Hématologie, Lille, France.

†† CHG, Centre de Traitement de l'Hémophilie, CHAMBERY, France

**Keywords:** DNA Mutational Analysis, FIX, Hemophilia B, Intellectual disability, Microarray

### 3. Résumé

**Contexte:** Plusieurs grandes délétions comprenant à la fois le gène *F9* dans sa totalité et d'autres gènes voisins ont été décrites chez des patients hémophiles B sévères. Certains de ces patients présentent des phénotypes cliniques atypiques associant une symptomatologie hémorragique et d'autres signes cliniques dont des déficits mentaux plus ou moins importants.

**Matériels et méthodes:** Dans cette étude, nous avons caractérisé par analyse chromosomique sur puce à ADN (CMA) six délétions totales du gène *F9* issues de la cohorte lyonnaise de familles d'hémophiles B. L'objectif de ce travail est d'étudier les corrélations génotype/phénotype chez ces patients.

**Résultats et discussion:** Un déficit mental était présent chez cinq des six patients inclus dans ce travail. L'analyse par CMA des génomes de ces patients a mis en évidence dans chaque cas la présence de délétions de très grande taille (de 933 kb à 9,19 Mb) situées au niveau des bandes Xq26.3 à Xq28. Dans tous les cas, la délétion totale du gène *F9* était associée à la perte d'un nombre variable de gènes voisins (de 5 à 28 autres gènes). L'analyse de nos données nous a permis d'associer à la présence d'un déficit mental à une zone de 1,26 Mb contenant sept gènes : *LOC389895*, *SOX3*, *LINC00632*, *CDR1*, *SPANXF1*, *LDOC1*, *SPANXC*. Le gène *SOX3* est le gène candidat retenu pour le déficit mental présent chez les 5 patients de notre cohorte. En effet, il code un facteur de transcription impliqué dans le développement du système nerveux central. A ce jour, tous les patients ayant une délétion du gène *SOX3* rapportés à la fois dans la littérature et dans cette étude présentent un déficit mental plus ou moins important.

**Conclusion:** Ce travail montre l'intérêt de réaliser une étude plus complète du génome chez des patients hémophiles B présentant un phénotype clinique atypique, notamment s'ils sont porteurs d'une délétion totale du gène *F9*. En effet la perte de gènes voisins du gène *F9*, tel que *SOX3*, peut avoir des conséquences cliniques importantes. La caractérisation précise de ces délétions par CMA prend toute son importance dans le cadre du conseil génétique donné aux femmes conductrices de ces anomalies afin d'évaluer le risque de transmission d'hémophilie B sévère associée à un déficit mental.

#### 4. Introduction

Le gène *F9* est situé sur le bras long du chromosome X (Xq27.1) et s'étend sur 32,7 kb. Il est constitué de 8 exons et est transcrit en un ARNm de 2802 pb codant une protéine de 461 acides aminés (135). Un grand nombre de mutations du gène *F9* responsables d'HB (> 1000 mutations répertoriées dans la base de données internationale, <http://www.factorix.org>) a été décrit. Parmi ces anomalies génétiques, les mutations ponctuelles sont les plus fréquentes (73%) suivies par les délétions (16%) et les insertions (3%) (136). Les délétions de grande taille (> 50pb) sont retrouvées dans environ 2% des cas d'HB sévère et leur présence est significativement corrélées à l'apparition d'inhibiteurs anti-FIX (13–15).

La délétion d'un ou de plusieurs exons du gène *F9* est généralement suspectée chez des patients hémophiles B sévères après plusieurs échecs d'amplification de ces exons par PCR. Cependant, cette approche n'est pas réalisable chez les femmes conductrices du fait de la présence de la copie du gène *F9* intacte sur le second chromosome X. Il existe néanmoins plusieurs méthodes pour étudier les variations du nombre de copie (CNVs) telles que la MLPA (multiplex ligation-dependent probe amplification) et la CMA. Une centaine de délétions de grande taille du gène *F9* ont été rapportées dans la littérature et 50% d'entre elles sont des délétions complètes du gène *F9*. Les points de cassure exacts n'ont été décrits que pour 18 de ces délétions et leur taille varie de 1,1 à 884 kb (137–148). Dans huit cas, la caractérisation précise des points de cassure a permis de mettre en évidence la délétion de gènes voisins du *F9* (145–148) et deux des patients porteurs de ces délétions présentaient un phénotype atypique d'hémophilie B sévère associant un syndrome hémorragique à d'autres signes cliniques dont un déficit mental.

Six patients porteurs d'une délétion totale du gène *F9* ont été étudiés au sein du laboratoire de génétique des pathologies hémorragiques des Hospices Civils de Lyon. Un déficit mental est présent chez cinq d'entre eux. L'objectif de ce travail a été de délimiter le plus précisément possible ces délétions sur le chromosome X par CMA afin d'étudier les corrélations génotype/phénotype chez ces patients.

## 5. Matériels et Méthodes

### Description de la population de l'étude

Nous avons étudié tous les patients (n = 6 ; 5 hommes et 1 femme) porteurs d'une délétion totale du gène *F9* issus de notre cohorte de 572 familles d'HB (dont 49% d'HB mineures, 13% de modérées et 28% de sévères). Les patients et leur famille ont été inclus dans cette étude après avoir signé un consentement éclairé. La délétion complète du gène *F9* a été suspectée suite à l'échec de l'amplification de tous les exons du gène chez les 5 patients hémophiles B sévères (P1 à P4 et P6). Ces patients ont été diagnostiqués durant les premières années de vie suite à des épisodes répétés de saignements prolongés. Des inhibiteurs anti-FIX ont été détectés chez les patients P1, P3 et P4.

Une délétion du gène *F9* à l'état hétérozygote a été mise en évidence par MLPA chez la patiente P5, une femme à taux bas de FIX (FIX:C = 10%) chez laquelle aucune autre mutation n'a été détectée après le séquençage de sa seconde copie du gène *F9*. La patiente P5 a été diagnostiquée de façon fortuite lors de la découverte d'un allongement isolé du TCA.

A l'exception du patient P6, tous les patients inclus dans ce travail présentaient un phénotype atypique associant un syndrome hémorragique et d'autres signes cliniques dont un déficit mental. Les patients ont été classés en trois catégories de déficit mental : sévère, modéré et mineur, en fonction de l'avis des cliniciens qui les prend en charge et de leur niveau scolaire (Tableau 8).



Patient	P1	P2	P3	P4	P5	P6
<b>Sexe</b>	Masculin	Masculin	Masculin	Masculin	Féminin	Masculin
<b>Age (ans)</b>	21	21	57	12	26	29
<b>FIX:C (%)</b>	<1	<1	<1	<1	10	<1
<b>Présence d'inhibiteurs</b>	Oui	Non	Oui	Oui	Non	Non
<b>Déficit mental</b>	Sévère	Mineur	Sévère	Mineur	modéré	Non
<b>Niveau de scolarisation</b>	Etablissement spécialisé	Etablissement spécialisé	Etablissement spécialisé	Scolarité normale	Etablissement spécialisé	Scolarité normale
<b>Retard du développement</b>	Aucun	Retard de la marche	Aucun	Retard de la marche et de la parole	Aucun	Aucun
<b>Autres manifestations cliniques</b>	Petite taille (-4SD à 16 ans)	Hypotonie, surpoids, cryptorchidie bilatérale ayant nécessité une orchidopexie, scoliose	Aucun	Hypotonie	Hypertrophie mammaire, cyphose	Aucun
<b>Déficit mental dans la famille</b>	Sœur (FIX:C =13%, DM mineur), Cousin maternel (FIX:C<1%, DM severe).	Aucun (Pas d'autre homophile B sévère dans la famille)	Frère (FIX:C<1%, inhibiteurs, DM severe), Cousin maternel (FIX:C<1%, DM severe)	Sœur (FIX:C = 20%, DM mineur) (Pas d'autre homophile B sévère dans la famille)	Aucun (Pas d'autre homophile B sévère dans la famille)	Aucun

**Tableau 8: Caractéristiques cliniques des patients inclus dans cette étude.** DM : Déficit mental, HB : hémophilie B.

### Réalisation des analyses chromosomiques sur puce à ADN (CMA)

Les CMA ont été réalisées par le laboratoire de cytogénétique des Hospices Civils de Lyon par Marie-Laure Carage dans le cadre de son mémoire de DES biologie médicale (65). La CMA a été réalisée sur de l'ADN génomique préalablement digéré et marqué par un fluorochrome (Alexa 3 ou Alexa 5). Une puce composée de 180 000 sondes, enrichies en oligonucléotides sur le chromosome X a été utilisée (SurePrint G3 Human CGH Microarray Kit, 4x180K, Agilent Technologies, Santa Clara, California, Etats-Unis). Au total, 100 000 sondes étaient situées sur le chromosome X et les 80 000 autres sondes étaient réparties sur le reste du génome. Le génome entier était couvert avec en moyenne une sonde tous les 120 kb. L'espacement médian entre deux sondes sur le chromosome X était de 3 kb. La couverture sur le gène *F9* est d'en moyenne une sonde tous les 617 pb. La CMA a été réalisée en suivant les recommandations du fournisseur. Les données ont été analysées à l'aide du logiciel Genomic Workbench 5.0.14. Une anomalie du nombre de copies dans une région du génome a été définie comme trois sondes successives ayant un ratio log2 anormal.

## 6. Résultats et discussion

Dans cette étude, six délétions complètes du gène *F9* ont été caractérisées par CMA afin d'étudier la relation entre la présence d'une large délétion dans cette région du chromosome X et l'existence d'un déficit mental. L'analyse cytogénétique a mis en évidence chez chacun des patients l'existence de délétions dans la région Xq26.3 à Xq28 de taille variable allant de 933 kb (P6) à 9,19 Mb (P5). Ces délétions ont été retrouvées à l'état hétérozygote chez la mère de chacun des patients à l'exception de la patiente P5 pour laquelle aucun des deux parents n'est porteur de l'anomalie génétique. L'apparition *de novo* de cette délétion chez la patiente P5 ou l'existence d'une mosaïque germinale chez la mère de cette patiente sont deux explications possibles de cette observation.

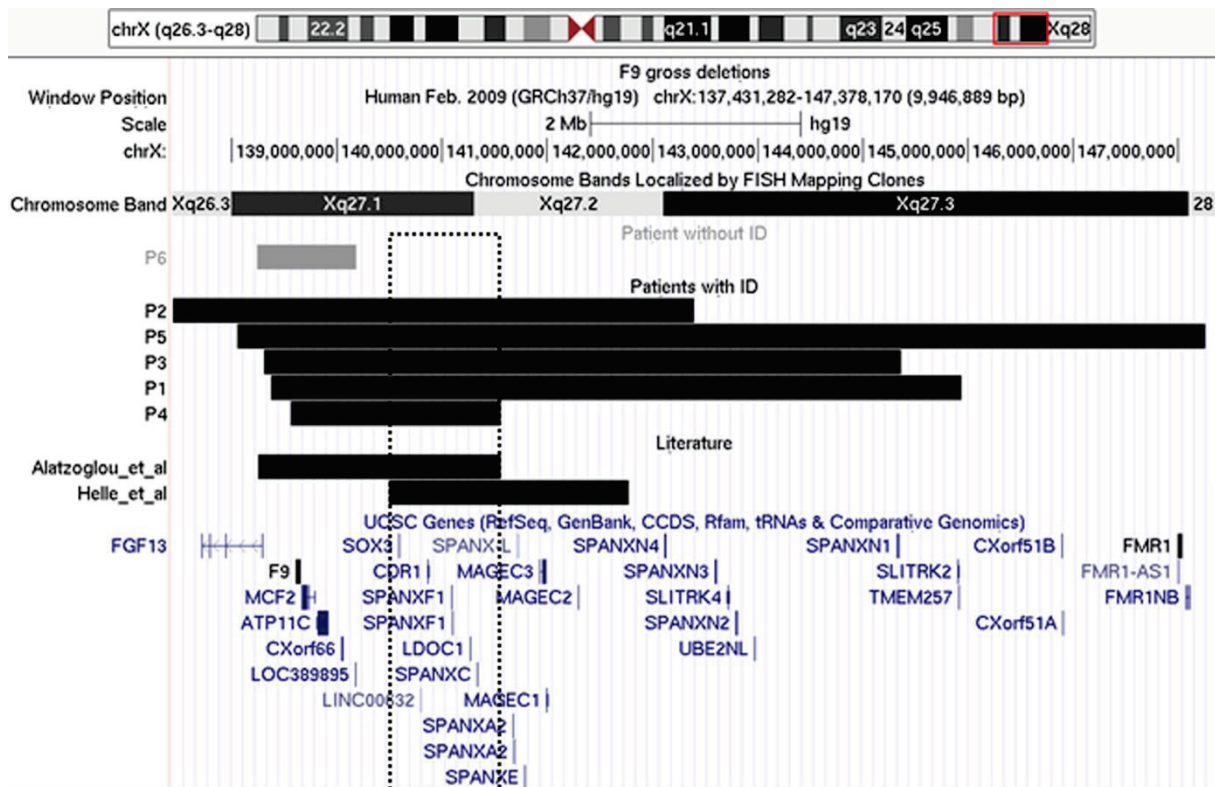
Chacune des six délétions étudiées est responsable de la perte du gène *F9* dans sa totalité ainsi qu'un nombre variable de gènes voisins (de 4 autres gènes délétés pour le patients P6 à 27 pour la patiente P5). Les caractéristiques de chacune des six délétions ont été reportées dans le Tableau 9. Comme ceci a été le cas dans d'autres travaux portant sur l'étude des grandes délétions du gène *F9*, aucun point de cassure récurrent n'a été mis en évidence (137–148).

Cinq des six patients inclus dans ce travail présentent un déficit mental plus ou moins important (P1-P5). La plus petite région de chevauchement des 5 délétions identifiées chez ces patients correspond à la région chrX: 138,383,974 – 140,443,615, hg19 (délétion du patient P4). Cette zone contient 11 gènes (*F9*, *MCF2*, *ATP11C*, *CXorf66*, *LOC389895*, *SOX3*, *LINC00632*, *CDR1*, *SPANXF1*, *LDOCI*, *SPANXC*). Le patient 6 qui ne présente pas de déficit mental est porteur de la plus petite délétion (chrX: 138,246,006 – 139,178,596, hg19) comprenant 5 gènes (*FGF13*, *F9*, *MCF2*, *ATP11C*, *CXorf66*). En faisant l'hypothèse d'une pénétrance complète, nous avons exclu de la liste des gènes candidats pour le déficit mental les gènes *F9*, *MCF2*, *ATP11C*, *CXorf66*. La plus petite région du génome associée au déficit mental a donc été réduite à une zone de 1,26 Mb (chrX: 139,178,596-140,443,615, hg19) contenant 7 gènes (*LOC389895*, *SOX3*, *LINC00632*, *CDR1*, *SPANXF1*, *LDOCI*, *SPANXC*). Les positions dans le génome de chacune de ces délétions sont présentées dans la Figure 12. L'HB et le déficit mental co-ségrègent chez les hommes atteints dans les familles des patients P1 et P3 (il n'y a aucun autre homme atteint dans les familles des patients P2, P4 et P5).

Patients	Coordonnées génomiques (taille maximum)	Taille (kb)	Gènes compris dans la délétion
P1	chrX: 138,197,325 – 144,743,347	6546.0	<i>F9, MCF2, ATP11C, CXorf66, LOC389895, SOX3, LINC00632, CDRI, SPANXF1, LDOC1, SPANXC, SPANXA2, SPANXL, SPANXE, MAGEC3, MAGEC1, MAGEC2, SPANXN4, SPANXN3, SLITRK4, SPANXN2, UBE2NL, SPANXN1, SLITRK2, TMEM257</i>
P2	chrX: 137,258,075 – 142,225,944	4967.9	<i>FGF13, F9, MCF2, ATP11C, CXorf66, LOC389895, SOX3, LINC00632, CDRI, SPANXF1, LDOC1, SPANXC, SPANXA2, SPANXL, SPANXE, MAGEC3, MAGEC1, MAGEC2, SPANXN4</i>
P3	chrX: 138,126,120 – 144,174,201	6033.5	<i>F9, MCF2, ATP11C, CXorf66, LOC389895, SOX3, LINC00632, CDRI, SPANXF1, LDOC1, SPANXC, SPANXA2, SPANXL, SPANXE, MAGEC3, MAGEC1, MAGEC2, SPANXN4, SPANXN3, SLITRK4, SPANXN2, UBE2NL, SPANXN1</i>
P4	chrX: 138,383,974 – 140,443,615	2059.6	<i>F9, MCF2, ATP11C, CXorf66, LOC389895, SOX3, LINC00632, CDRI, SPANXF1, LDOC1, SPANXC</i>
P5	chrX: 137,880,627 – 137,884,787	9191,6	<i>F9, MCF2, ATP11C, CXorf66, LOC389895, SOX3, LINC00632, CDRI, SPANXF1, LDOC1, SPANXC, SPANXA2, SPANXL, SPANXE, MAGEC3, MAGEC1, MAGEC2, SPANXN4, SPANXN3, SLITRK4, SPANXN2, UBE2NL, SPANXN1, SLITRK2, TMEM257, CXorf51A, CXorf51B, FMRI</i>
P6	chrX: 138,246,006 – 139,178,596	932.5	<i>FGF13 (partiellement), F9, MCF2, ATP11C, CXorf66,</i>

**Tableau 9: Caractérisation de 6 délétions complètes du gène *F9* par CMA**

L'association de l'HB sévère et d'un déficit mental a été décrite pour la première fois chez un homme ayant une délétion de 6Mb sur le chromosome X (Xq26.1-27.2) (149). Depuis, deux autres cas d'association HB et déficit mental ont été rapportés dans littérature : une délétion d'au moins 2Mb chez un homme autiste, épileptique et présentant un retard du développement et une délétion de 2,61 Mb dans la région Xq27.1-27.2 trouvée chez un homme présentant un retard du développement et une insuffisance hypophysaire (146,147). Nous avons recherché dans la littérature si d'autres délétions dans la région Xq27.1 avaient été rapportées et nous avons trouvé la description d'une délétion de 2,1 Mb (chrX: 139,498,435-141,770,365, hg19) chez un jeune garçon de 8 ans présentant un déficit mental léger et des troubles du langage et du comportement (150). Les points de cassure de ces deux dernières délétions ont été déterminés par CMA (147,150). Ainsi en prenant en compte ces données de la littérature, nous avons réduit notre région candidate pour le déficit mental à une zone de 1,06 Mb (chrX: 139,498,435-140,557,617, hg19) contenant 6 gènes (*SOX3, LINC00632, CDRI, SPANXF1, LDOC1, SPANXC*) (Figure 12).



**Figure 12 : Coordonnées génomiques des différentes délétions étudiées (UCSC, hg19).** Les délétions mises en évidence chez les six patients inclus dans ce travail (P1-P6) ainsi que deux délétions de la région Xq27 rapportées dans la littérature sont présentées. Les délétions des patients présentant un déficit mental sont représentées en noir et celle du patient ne présentant pas de déficit mental est représentée en gris. La région de 1,06 Mb liée à la présence d'un déficit mental est indiquée par des pointillés.

Le gène *CDR1* (Cerebellar Degeneration-Related autoantigen 1) code une protéine de fonction inconnue. Les gènes *SPANXF1* et *SPANXC* (Sperm Proteins Associated with the Nucleus) appartiennent à une famille de gènes codant des protéines impliquées dans la spermiogénèse. *LDOC1* (Leucine zipper DOWNregulated in Cancer 1) est un gène suppresseur de tumeur (151). *LINC00632* est une séquence non codante (elle est transcrite en un ARN non traduit). En prenant en compte les connaissances actuelles sur les fonctions de ces gènes et le fait qu'aucun d'entre eux ne soit impliqué dans le développement des structures neurologiques, il nous a semblé peu probable que l'un de ces 5 gènes soit impliqué dans le déficit mental observé chez nos patients.

A l'inverse, une délétion du gène *SOX3* pourrait en partie être responsable du déficit mental présent chez les patients P1 à P5. En effet, *SOX3* code un facteur de transcription impliqué dans le développement de l'hypophyse (147). La protéine *SOX3* appartient à la famille des facteurs de transcription SOX (SRY-related HMG box) (152). Cette protéine est composée d'un domaine N-

terminal de fonction inconnue, d'un site de liaison à l'ADN de type HMG (High-Mobility Group) et d'un long domaine C-terminal d'activation de transcription composée de 4 régions contenant des répétitions d'alanines. Des expériences ont montré la présence de déficit des fonctions hypophysaires et des anomalies du système nerveux central chez la souris KO pour le gène *SOX3* (153). Chez l'homme, des mutations de la région C-terminale (expansions d'alanines ou délétions) et les duplications du gène *SOX3* sont associées à des phénotypes variables de déficits hypophysaires liés à l'X allant du déficit isolé en hormone de croissance au panhypopituitarisme associés dans certains cas à un déficit mental (147). Par contre, la totalité des patients de la littérature ayant une délétion totale du gène *SOX3* présentent un déficit mental (146,147,150,152).

Deux autres gènes, *FGF13* (Fibroblaste Growth Factor 13) et *FMRI* (Fragile X Mental Retardation 1) sont impliqués dans des déficits mentaux liés à l'X. Le gène *FGF13* est délété chez les patients P2 et P6 et les gènes *FGF13* et *FMRI* sont délétés chez la patiente P5. *FGF13*, un gène codant une protéine appartenant à la famille FGF, est exprimé dans les neurones du cortex cérébral durant le développement embryonnaire. Il a été décrit comme un gène candidat potentiel pour certains déficits mentaux syndromiques liés à l'X. En effet, il a été montré l'existence de déficit de la migration des neurones dans le cortex et l'hippocampe chez des souris KO pour ce gène ainsi que des difficultés d'apprentissage et des troubles de la mémoire (154). Cependant, une délétion du gène *FGF13* a été trouvée chez 3 de nos patients dont le patient P6 qui ne présente pas de déficit mental. Le gène *FMRI* est associé aux syndromes de l'X fragile et de tremblement-ataxie lié à une prémutation de l'X fragile (FXTAS). Chacune de ces deux formes cliniques est caractérisée par de nombreux symptômes et notamment un déficit mental (155). Ainsi, il est possible que la délétion partielle hétérozygote du gène *FMRI* soit en partie responsable du déficit mental observé chez la patiente P5.

Dans notre travail, aucune femme porteuse d'une délétion hétérozygote du gène *F9* ne présentait un déficit mental sévère. Il est néanmoins intéressant de noter que la patiente P5 et les sœurs des patients P1 et P4 présentent un déficit mental mineur à modéré. Ces trois patientes sont des femmes à taux bas de FIX (FIX:C : 10-20%). Il est connu que les femmes porteuses de mutations du gène *F9* à l'état hétérozygote peuvent avoir des taux de FIX et des phénotypes hémorragiques

variables (156,157). Ceci est attribué en partie au phénomène d'inactivation de l'X décrit par Lyon *et al* en 1962 (158). Le phénomène de lyonisation de l'X pourrait aussi expliquer la présence d'un déficit mental chez nos trois femmes à taux bas de FIX. Nous ne disposons pas des dosages de FIX chez la totalité des femmes conductrices de ce travail, il est donc impossible de conclure sur le lien entre présence d'un déficit mental et un taux bas de FIX chez ces patientes.

En conclusion, dans ce travail, nous avons caractérisé par CMA la plus grande cohorte de délétions totales du gène *F9* rapportée à ce jour. Nos résultats suggèrent fortement que la délétion du gène *SOX3* soit associée à la présence d'un déficit mental plus ou moins important chez les patients. Ainsi, les patients hémophiles B présentant des phénotypes cliniques atypiques, notamment ceux porteurs d'une délétion totale du gène *F9* à l'état hémizygote doivent être orientés vers un généticien afin d'évaluer l'intérêt de la réalisation d'une étude plus globale de leur génome. De telles investigations sont particulièrement importantes chez les femmes conductrices de délétions totales du gène *F9* afin de déterminer le risque de transmission de l'hémophilie B associée à un déficit mental.

**Partie 2. Caractérisation par analyse chromosomique sur puce à ADN de cinq réarrangements génomiques complexes affectant le gène *F8* : intérêt pour le conseil génétique.**

**Molecular cytogenetic characterization of five F8 complex rearrangements: utility for haemophilia A genetic counselling.**

Haemophilia. 2017, 23: e316-e323.

**Y. Jourdy**,\*†# N. Chatron,‡§# M. Fretigny,\* M. L. Carage,\* H. Chambost,¶  
S. Claeysens-Donadel,\*\* V. Roussel-Robert,†† c. Negrier,\*† D. Sanlaville‡§ and C. Vinciguerra\*†

\*Hospices Civils de Lyon, Hôpital Edouard Herriot, Service d'hématologie Biologique;

†EA 4609 Hémostase et cancer, Université Claude Bernard Lyon 1, Lyon;

‡Hospices Civils de Lyon, Groupe Hospitalier Est, Laboratoire de Cytogénétique Constitutionnelle, Bron, France;

§CRNL, équipe GENDEV INSERM U1028, CNRS UMR5292, Université Claude Bernard Lyon 1, Lyon;

¶Centre de traitement de l'hémophilie, CHU La Timone, Marseille;

\*\*Centre de traitement de l'hémophilie, CHU Purpan, Toulouse;

††Centre de traitement de l'hémophilie, Hôpital Cochin, Paris, France

#Co-auteurs

**Keywords:** Genomic rearrangement, Haemophilia A, Microarray Analysis, Intron 22 inversion, Intron 1 inversion.

## 1. Résumé

**Contexte :** Les inversions génomiques sont classiquement équilibrées mais des profils anormaux d'inversion des introns 1 et 22 ont été décrits chez des patients hémophiles A sévères suggérant l'existence de réarrangements plus complexes associant inversions et délétions et/ou duplications.

**Objectifs et méthodes :** Quatre profils anormaux d'inversion de l'intron 22 et un profil anormal d'inversion de l'intron 1 ont été étudiés par analyse chromosomique sur puce à ADN L'objectif de ce travail est d'étudier les corrélations génotype/phénotype chez ces patients.

**Résultats et discussion :** L'analyse cytogénétique de chaque inversion a montré l'existence de réarrangements complexes de la région Xq28. Chez trois patients, des duplications de grande taille, allant de 230 kb à 1302 kb, ont été mises en évidence. Ces duplications englobent de nombreux gènes voisins du gène *F8* dont le gène *RAB39B*. La duplication du gène *RAB39B* est une anomalie génétique candidate pour certains déficits mentaux liés à l'X. Néanmoins, aucun des patients hémophiles A sévères porteurs d'une duplication du gène *RAB39B* décrits à ce jour (à la fois dans la littérature et dans cette étude) ne présente de déficit mental. L'une des hypothèses possibles à cette observation est que les réarrangements complexes affectant le gène *F8* pourraient entraîner une régulation négative de l'expression de *RAB39B*. Chez les deux derniers patients, l'analyse cytogénétique a montré l'existence de délétions de grande taille (de 285 à 522 kb). Un syndrome Moyamoya est fortement suspecté chez un de ces deux patients porteur d'une délétion du gène *BRCC3*.

**Conclusion :** Plusieurs gènes voisins du *F8* sont associés à différentes pathologies telles que des déficits mentaux ou des maladies cardiovasculaires. Tous les patients hémophiles A chez lesquels un réarrangement complexe de la région Xq28 est suspecté doivent donc être orientés vers un généticien afin d'évaluer l'intérêt de la réalisation d'une étude plus globale de leur génome. De telles investigations sont particulièrement importantes chez les femmes conductrices de ces anomalies afin de déterminer le risque de transmission de l'hémophilie A sévère associée à un « syndrome des gènes contigus ».



## 2. Introduction

Une grande variété d'anomalies génétiques affectant le gène *F8* peut être responsable d'HA. Des inversions génomiques de grande taille connues sous le nom d'inversion de l'intron 22 (Inv22) et d'inversion de l'intron 1 (Inv1) sont responsables respectivement de 43% et 2,3% des HA sévères (56–59). Les inversions génomiques sont classiquement équilibrées mais des profils anormaux d'Inv22 et Inv1 ont été rapportés dans la littérature suggérant l'existence de mécanismes moléculaires plus complexes associant inversions et duplications ou délétions (159–169). Dans cette étude nous rapportons des réarrangements génomiques non équilibrés mis en évidence chez 4 patients hémophiles A sévères et une femme conductrice à taux bas de FVIII. Chacun de ces patients présentait un profil anormal soit en Southern blot (SB) pour l'Inv22 soit en PCR pour l'Inv1. Une analyse par CMA a été réalisée chez ces patients afin de caractériser plus précisément ces réarrangements génomiques.

## 3. Matériels et méthodes

### Patients

Parmi notre cohorte de 1879 familles d'hémophiles A (phénotype mineur, n=921 ; modéré, n=262 ; sévère, n=696), 7 patients présentaient des profils anormaux soit en SB pour l'Inv22 (n=5/7) ou bien en PCR pour l'Inv1 (n=2/7). Cinq de ces 7 patients pour lesquels un échantillon d'ADN était disponible ont été inclus dans ce travail (patients 1-5). Tous les patients ont signé un consentement éclairé pour la réalisation d'analyses génétiques.

### Détection de l'inversion de l'intron 1

La méthode utilisée pour la détection de l'Inv1 est celle décrite par Bagnall *et al.* (67) et consiste en la réalisation de deux PCR. Lors de la première PCR, des amorces spécifiques de la région *int1h-1* (9F et 9CR) ainsi qu'une amorce se fixant spécifiquement du côté centromérique d'*int1h-2* (h2F) ont été utilisées. Lors de la 2<sup>nde</sup> réaction, des amorces spécifiques de la région *int1h-2* (H2F et 2R) ainsi qu'une amorce se fixant spécifiquement du côté télomérique d'*int1h-1* (9F) ont été utilisées. Les séquences des amorces utilisées, la composition des mélanges réactionnels et les conditions de réactions sont décrites dans les Tableau 10, Tableau 11 et Tableau 12.

Amorces	Séquence 5'-3'
9F	GTGTTGGGAATGGTTACGG
9CR	CTAGCTTGAGCTCCCTGTGG
H2F	GGCAGGGATCTTGTGGTAAA
2R	TGGGTGATATAAGCTGCTGAGCTA

**Tableau 10 : Séquences des amorces utilisées pour les deux PCR nécessaires à la mise en évidence d'une inversion de l'intron 1.**

Réactifs	Volume (µl)	
	PCR 1	PCR2
Tampon I AmpliTaq 10X	5	5
ADN patient ou témoin	200 ng	200 ng
Amorce 9F [10mM]	3	3
Amorce 9CR [10mM]	3	0
Amorce H2F [10mM]	3	3
Amorce 2R [10mM]	0	3
dNTP [2,5 mM de chaque dNTP]	4	4
AmpliTaq Polymerase [2,5U/µl]	0,5	0,5
Eau	Qsp 50	Qsp 50

**Tableau 11 : Composition des mélanges réactionnels pour les PCR nécessaires à la mise en évidence d'une inversion de l'intron 1.**

	Température (°C)	Temps	Nombre de cycles
Dénaturation initiale	94	2 min	
Dénaturation	94	30 s	
Hybridation	65	30 s	30
Elongation	72	2 min	
Elongation finale	72	7 min	

**Tableau 12 : Conditions de PCR pour la recherche de l'inversion de l'intron 1.**

La taille des produits obtenus avec ces deux PCR permet de déterminer la présence ou non d'une Inv1. Des produits de PCR de 1908 pb (1<sup>ère</sup> PCR) et de 1191 pb (2<sup>nde</sup> PCR) correspondent à un profil normal (non inversé) alors que des bandes de 1303 pb (1<sup>ère</sup> PCR) et 1776 pb (2<sup>nde</sup> PCR) indiquent la présence d'une Inv1.

## Détection de l'inversion de l'intron 22

La recherche de la présence d'une Inv22 a été réalisée à la fois par SB et LR-PCR chez tous les patients.

### Southern Blot

Le SB a été réalisé comme précédemment décrit (56). Dans une première étape, 10µg d'ADN génomiques sont digérés par l'enzyme de restriction *BclI* (50U) une nuit à 50°C. Les ADN digérés (40µl) sont ensuite migrés sur gel d'agarose 0,5% à 70V pendant 24h. Après deux rinçages avec de l'eau, le gel est traité 30 minutes avec une solution dénaturante (NaCl 3M, NaOH 0,5N) à température ambiante sous agitation. Après deux nouveaux rinçages, le gel est traité par une solution neutralisante (NaCl 1,5 M, Tris 0,5M, pH 7), 30 minutes à température ambiante sous agitation puis, après deux

lavages, le gel est traité 30 min par du tampon SSC 10X (tampon citrate sodique de sel) à température ambiante sous agitation. L'ADN est ensuite transféré sur une membrane de Nylon. Le lendemain du transfert, la membrane de nylon est séchée 2h à 80°C puis humidifiée dans du tampon SSC 2X. La membrane est ensuite trempée dans une solution de pré-hybridation (10,4 ml de solution de Na<sub>2</sub>HPO<sub>4</sub>·2H<sub>2</sub>O 1,2M pH 7,2, 8,8 ml de SDS 20%, 4,8ml d'eau) et incubée 4h à 60°C sous rotation. Pour l'étape d'hybridation, 10µl de sonde d'hybridation spécifique de la région *int22h* marquée à la digoxine sont dénaturés à 94°C pendant 10 minutes puis ajoutés à 5ml de solution d'hybridation (DIG Easy Hyb, Sigma Aldrich, Saint Louis, MI, EU). Le tampon de préhybridation est éliminé et remplacé par la solution d'hybridation. La membrane est incubée une nuit à 60°C sous rotation. Le lendemain, la membrane est lavée une première fois (tampon SSC 0,5X, SDS 0,1%) à 55°C pendant 20 minutes puis une seconde fois dans une solution de SSC 0,1X à 55°C pendant 20 minutes. Pour l'étape de révélation, la membrane est incubée 30 minutes à température ambiante en présence d'anticorps anti-digoxine marqués à la phosphatase alcaline, après plusieurs étapes de lavage, les bandes d'intérêt sont révélées par chimiluminescence. Le SB permet la détection des trois régions *int22h*. Des bandes de 21,5, 14 et 16 kb correspondent respectivement aux régions *int22h-1*, *2* et *3* non inversées. L'Inv22-1 est caractérisée par la présence des bandes de 20 (*int22h-3/1*), 14 (*int22h-2*) et 17,5kb (*int22h-1/3*) et l'Inv22-2 est définie par la présence des bandes de 20 (*int22h-2/1*), 15,5 (*int22h-1/2*) et 16kb (*int22h-3*).

#### PCR long fragment

Une LR-PCR utilisant 5 amorces a été utilisée pour confirmer les résultats du SB. Cette technique a été décrite par Bagnall *et al.* (63). La présence de 5 amorces permet l'amplification simultanée des 3 régions *int22h*. En fonction de la taille des produits de PCR obtenus il est possible de mettre en évidence la présence d'une Inv22 de type 1 ou 2. Les séquences des amorces utilisées, la composition des mélanges réactionnels et les conditions de réaction sont décrites dans les Tableaux 13, 14 et 15.

Amorces	Séquence 5'-3'
H1F	CCCTTACAGTTATTAACACTACTCTCATGAGGTTTCATTCC
H1R	CCCCGGCACTTGAAAGTAGCAGATGCAAGAAGGGCACA
H2F	TCTCAAATAAGGAATACAGAGAAACCTGAACAGACC
H3F	TGGCTGAATGAAACCAAGTGTGAGCCATAAATCA
H2/3R	ACTATAACCAGCACCTTGAACCTTCCCCTCTCATA

**Tableau 13 : Séquences des amorces de la PCR long fragment utilisées pour la mise en évidence d'une inversion de l'intron 22.**

Réactifs	Volume (µl)
Tampon I (expand long Template) 10X	2,5
ADN patient ou témoin	250 ng
MgCl <sub>2</sub> [25mM]	0,5
DMSO	1,875
dNTP (dATP [5mM], dCTP [5mM], dTTP [5mM], dGTP [2,5mM], 7-deaza-GTP [2,5mM])	2,5
Expand Long Template PCR System, Enzyme mix (5U/µl)	0,66
H1F [40mM]	0,18
H1R [20mM]	0,25
H2F [5mM]	1
H3F [12mM]	0,25
H2/3R [20mM]	0,5
Eau	Qsp 25

**Tableau 14 : Composition du mélange réactionnel pour la PCR long fragment utilisée pour la mise en évidence d'une inversion de l'intron 22.**

	Température (°C)	Temps	Nombre de cycles
Dénaturation initiale	94	2 min	
Dénaturation	94	12 s	
Hybridation / élongation	65	15 min	28
Elongation finale	68	12 min	

**Tableau 15 : Conditions de PCR pour la recherche de l'inversion de l'intron 22.**

Les produits de PCR sont migrés une nuit à 70V en gel d'agarose 0,8%. La LR-PCR permet la détection des trois régions *int22h*. Des bandes de 10, 12 et 11kb correspondent respectivement aux régions *int22h-1*, 2 et 3 non inversées. L'Inv22-1 est caractérisée par la présence des bandes de 9,6 (*int22h-3/1*), 12 (*int22h-2*) et 11,5 kb (*int22h-1/3*) et l'Inv22-2 est définie par la présence des bandes de 9,6 (*int22h-2/1*), 12,7 (*int22h-1/2*) et 11kb (*int22h-3*).

Les différents profils caractéristiques obtenus avec les deux méthodes (SB et LR-PCR) pour la détection de l'inversion de l'intron 22 sont présentés dans le Tableau 16.

Profil	Taille des bandes (kb) Southern Blot			Taille des bandes (kb) PCR Long fragment		
	<i>int22h-1</i>	<i>int22h-2</i>	<i>int22h-3</i>	<i>int22h-1</i>	<i>int22h-2</i>	<i>int22h-3</i>
Normal	21,5	14	16	10	12	11
Inv22-1 (homme)	20	14	17,5	9,6	12	11,5
Inv22-2 (homme)	20	15,5	16	9,6	12,7	11

**Tableau 16 : Résumé des profils typiques obtenus avec les méthodes SB et LR-PCR pour la mise en évidence des inversions de l'intron 22.**

## Réalisation des analyses chromosomiques sur puce à ADN (CMA)

Les analyses cytogénétiques ont été réalisées par le laboratoire de cytogénétique des Hospices Civils de Lyon par Marie-Laure Carage dans le cadre de son mémoire de DES biologie médicale (65). La méthode est présentée dans la partie 1 de cette section. La couverture sur le gène *F8* était d'en moyenne une sonde tous les 900 pb.

Afin de confirmer les résultats de CMA, les variations du nombre de copies (CNV) des exons du gène *F8* ont été contrôlées par MLPA en suivant les recommandations du fournisseur (MRC-Holland, Amsterdam, Pays-Bas).

### 4. Résultats

Parmi les cinq patients inclus dans cette étude, 4 sont des hommes hémophiles A sévères (patients 1 à 4) et la patiente 5 est une femme conductrice à taux bas de FVIII (FVIII:C = 10%). Seul le patient 3 a développé des inhibiteurs anti-FVIII. En plus des manifestations cliniques liées à l'HA sévère, le patient 3 est de petite taille (1,55 m à l'âge adulte), a eu une puberté précoce et présente une macrocéphalie et un faciès lunaire. Les caractéristiques cliniques des patients sont répertoriées dans le Tableau 17.

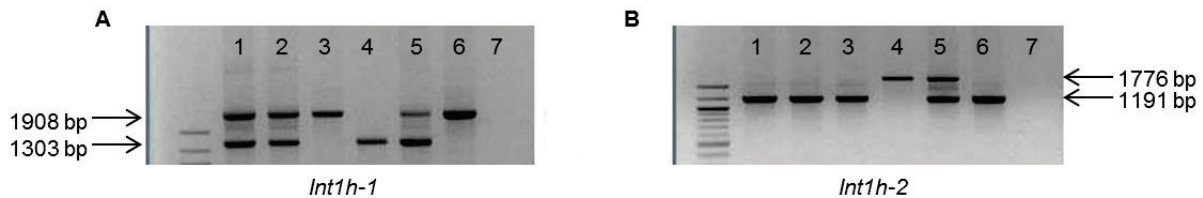
Patient	P1	P2	P3	P4	P5
Sexe	Homme	Homme	Homme	Homme	Femme
FVIII:C (%)	<1	<1	<1	<1	10
Présence d'inhibiteurs	Non	Non	Oui	Non	Non
Autres signes cliniques	Aucun	Aucun	Petite taille (1,55 m à l'âge adulte), puberté précoce, macrocéphalie et un faciès lunaire.	Aucun	Aucun

Tableau 17 : Caractéristiques cliniques des patients de l'étude.

#### Patient 1

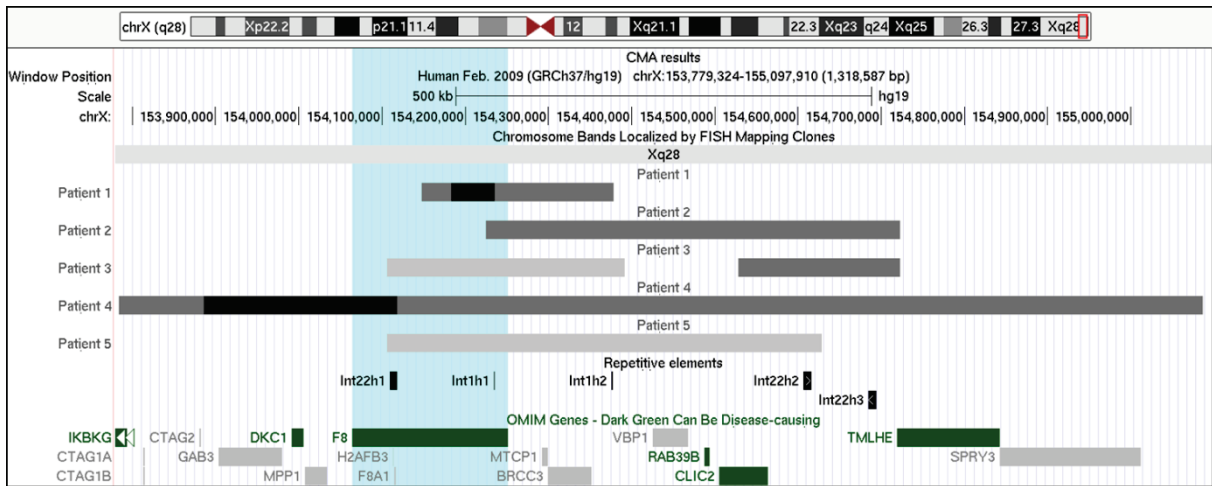
Les recherches des Inv1 et Inv22 ont été réalisées chez le patient 1. Les résultats de SB et LR-PCR ont permis d'exclure une Inv22 mais un profil anormal de PCR de l'Inv1 a été mis en évidence avec la présence d'une bande surnuméraire. La taille de la première bande correspond à la taille de la séquence *int1h-1* normale (1908 pb), celle de la seconde bande correspond à la taille de la séquence hybride *int1h-1/2* (1303 pb) indiquant la présence d'une Inv1 et la taille de la troisième bande

correspond à la taille de la séquence *int1h-2* normale (1191 pb) (Figure 13). Ces résultats suggèrent la présence de copies surnuméraire des régions *int1h-1* et *int1h2*. L'absence de la bande correspondant à l'hybride *int1h2/1* pourrait être due à un réarrangement de la région flanquant cette séquence empêchant son amplification par PCR.



**Figure 13 : Profil anormal de l'inversion de l'intron 1.** A : résultat de l'amplification de la région *int1h-1*. 1 et 2 : patient 1 ; 3 et 6 : homme sain ; 4 : homme avec une inversion de l'intron 1 ; 5 femme conductrice d'une inversion de l'intron 1 ; 7 : contrôle négatif. B : résultat de l'amplification de la région *int1h-2*. 1 et 2 : patient 1 ; 3 et 6 : homme sain ; 4 : homme avec une inversion de l'intron 1 ; 5 femme conductrice d'une inversion de l'intron 1 ; 7 : contrôle négatif. Le contrôle négatif correspond à une PCR sans ADN.

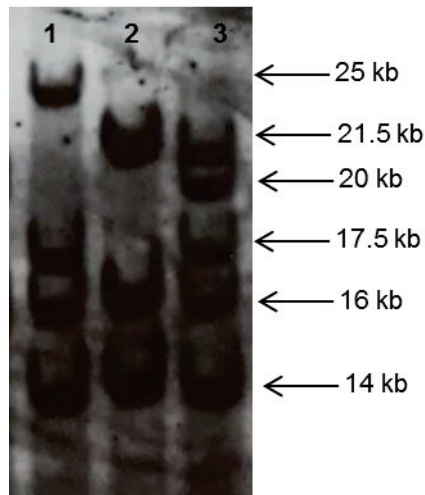
L'analyse par CMA du génome du patient 1 a mis en évidence l'existence d'un réarrangement génétique complexe de grande taille avec la présence d'une duplication de 230 kb (chrX:154,147,182-154,378,045, hg19) débutant dans l'intron 14 du gène *F8* et s'étendant jusqu'au gène *BRCC3*. Cette duplication comprend 5 gènes: *F8*, *FUNDC2*, *CMC4*, *MTCPI* et *BRCC3* ainsi que *int1h-2* qui est localisé 24kb en aval du gène *BRCC3*. Cette duplication est combinée à une triplification de 52,5 kb s'étendant des exons 1 à 11 du gène *F8* (chrX:154,182,749-154,235,426, hg19). La duplication des exons 12 à 14 et la triplification des exons 1 à 11 du gène *F8* ont été confirmés par MLPA (Figure 14 et Tableau 18).



**Figure 14 : Coordonnées génomiques des différents réarrangements étudiés (UCSC, hg19).** Les régions dupliquées (gris foncé), tripliquées (noire) et délétées (gris claire) trouvées chez les 5 patients sont indiquées sur la figure. La section « repetitive elements » indique les positions des séquences répétées *int1h* et *int22h*. Les gènes en vert sont associés à des pathologies connues.

## Patient 2

Les résultats du SB ont montré la présence de 4 bandes (14, 16, 17,5 et 25 kb). Les bandes de 14 et 17,5kb (correspondant respectivement à *int22h-2* et à l'hybride *int22h-1/3*) indiquent l'existence d'une *Inv22-1*. Cependant la taille de la troisième bande est clairement supérieure aux 20 kb attendus pour la séquence *int22h-3/1* (approximativement 25 kb) et la bande de 16 kb (correspondant à la séquence *int22h-3* non inversée) est normalement absente de ce type d'inversion (Figure 15). Les résultats de LR-PCR ont confirmé le SB. Quatre bandes ont été observées (10, 11, 11,5 et 12 kb) dont deux caractéristiques d'une *Inv22-1*, une troisième bande plus grande que la bande attendue (10 vs 9,6 kb) et une bande de 11 kb (*int22h-3* non inversée) normalement absente dans le cas d'une *Inv22-1*. Ces résultats sont en faveur de la présence de copies surnuméraires d'*int22h-3* et d'un réarrangement de la région flanquant *int22h-1* responsable de la présence d'une bande plus grande en SB et LR-PCR.



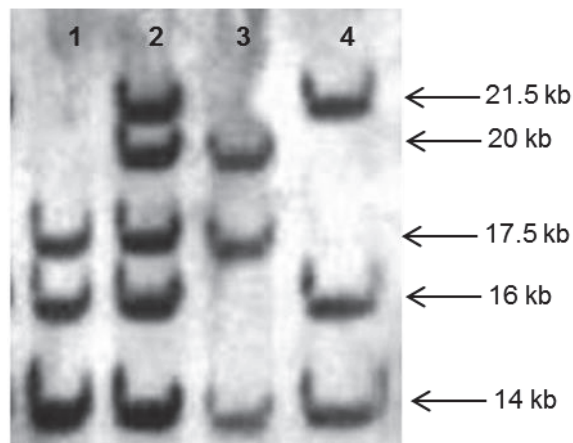
**Figure 15 : Résultats de Southern blot pour le patient 2.** 1 : patient 2 ; 2 : homme sain ; 3 : femme conductrice d'une inversion de l'intron 22 de type 1.

L'analyse par CMA du génome du patient 2 a trouvé une duplication de 498 kb (chrX:154,224,460-154,722,370, hg19) englobant 12 gènes et les 3 premiers exons du gène *F8* (Figure 14 et Tableau 18). La MLPA a confirmé la duplication des exons 1 à 3 du gène *F8*.

### Patient 3

Les résultats du SB ont montré la présence de 3 bandes de même intensité (14, 16 et 17,5kb). Les bandes de 14 et 17,5kb (correspondant respectivement à *int22h-2* et à l'hybride *int22h-1/3*) indiquent l'existence d'une Inv22-1. Cependant la bande de 20 kb (hybride *int22h-3/1*) caractéristique d'une Inv22-1 est absente et la bande de 16 kb (correspondant à la séquence *int22h-3* non inversée) est normalement absente de ce type d'inversion (Figure 16). Les résultats de LR-PCR ont confirmé le SB. Trois bandes ont été observées (12, 11 et 11,5 kb) dont deux caractéristiques d'une Inv22-1 et une bande de 11 kb (*int22h-3* non inversée) normalement absente dans le cas d'une Inv22-1. Le fragment de 9,6 kb (hybride *int22h-3/1*) est absent. Ces résultats sont en faveur de la présence de copies surnuméraires d'*int22h-3* et d'une délétion de la région comprenant la séquence hybride *int22h-3/1*.



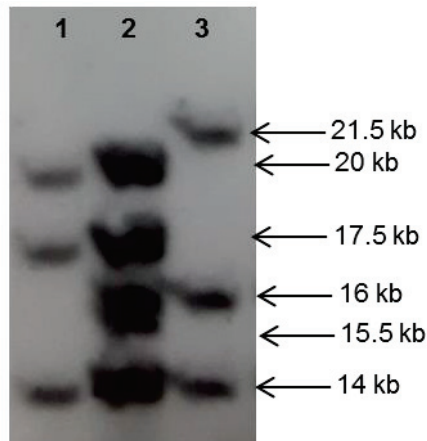


**Figure 16 : Résultats de Southern blot pour le patient 3.** 1 : patient 3 ; 2 : femme conductrice d'une inversion de l'intron 22 de type 1 ; 3 : homme porteur d'une inversion de l'intron 22 de type 1, 4 : homme sain.

L'analyse par CMA du génome du patient 3 a montré la présence d'une délétion de 285 kb (chrX: 154,106,243-154,391,634, hg19) comprenant 4 gènes (*FUNDC2*, *CMC4*, *MTCP1* et *BRCC3*) ainsi que les exons 1 à 22 du gène *F8* (confirmé par MLPA). Cette délétion est associée à une duplication de 194 kb (chrX: 154,528,419-154,722,429, hg19) allant de l'intron 2 du gène *CLIC2* à la séquence *int22h-3* (Figure 14 et Tableau 18). Cette délétion permet d'expliquer la présence de la bande supplémentaire *int22h-3* sur les résultats du SB et de la LR-PCR. La mère et la sœur du patient 3 sont porteuses de la même anomalie génétique.

#### **Patient 4 :**

Les résultats du SB ont montré la présence de 5 bandes (14, 15,5, 16, 17,5 et 20 kb). Ce résultat suggère la présence d'une duplication complète du gène *F8* associée à une Inv22-1 (14, 17,5, 20 kb) et une Inv22-2 (15,5, 16 et 20 kb ; Figure 17). La LR-PCR a confirmé ces résultats et a montré 5 bandes (9,6, 11, 11,5, 12 et 12,7 kb).



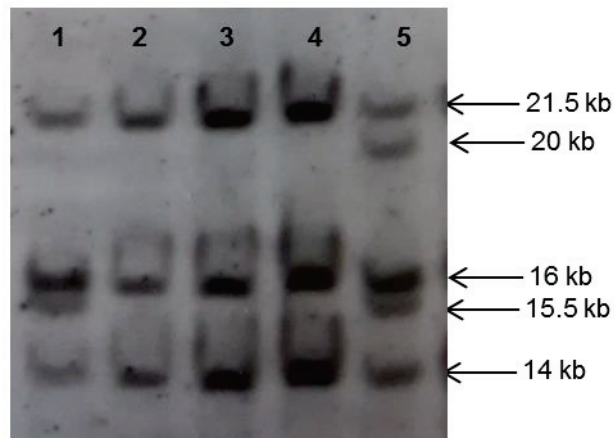
**Figure 17 : Résultats de Southern blot pour le patient 4.** 1 : homme porteur d'une inversion de l'intron 22 de type 1 ; 2 : patient 4 ; 3 : homme sain.

L'analyse par CMA du génome du patient 4 a montré la présence d'un réarrangement plus complexe avec une duplication de 1302 kb (chrX: 153,783,581-155,086,461, hg19) englobant 23 gènes associée à deux triplications : la première s'étendant du gène *GAB3* jusqu'à l'intron 22 du gène *F8* (233kb, ChrX:153,885,800-154,118,678, hg19) et la seconde allant de l'intron 2 du gène *CLIC2* à la séquence *int22h-3* (194 kb chrX: 154,528,419-154,722,429, hg19 ; Figure 14 et Tableau 18). La MLPA a confirmé la triplication des exons 22 à 26 et la duplication des exons 1 à 21 du gène *F8*. La mère et le frère (hémophile A sévère) du patient 4 sont porteurs de la même anomalie génétique.

### **Patiente 5**

La patiente 5 est une femme conductrice d'hémophilie A avec un taux bas de FVIII (FVIII:C = 10%). Elle est hétérozygote composite avec un profil anormal d'Inv22 sur une copie du gène *F8* et l'association de mutations *F8* c.2044G>C (p.Val682Leu) + *F8* c.2887G>A (p.Asp963Asn) sur l'autre copie du gène *F8*. Sa mère et sa grand-mère sont porteuses de l'association des deux mutations faux-sens. Les résultats du SB ont mis en évidence 4 bandes (14, 15,5, 16 et 21,5 kb) de même intensité à l'exception de la bande de 16 kb qui présente une intensité supérieure. Ce profil suggère la présence d'une Inv22-2 à l'état hétérozygote mais la bande de 20 kb caractéristique de cette inversion (*int22h-2/1*) est absente (Figure 18). Des résultats

similaires ont été retrouvés en LR-PCR : présence de 4 bandes (10, 11, 12, 12,7kb) et la bande correspondant à l'hybride *int22h-2/1* est manquante (9,6 kb).



**Figure 18 : Résultats de Southern blot pour la patiente 5.** 1 : patiente 5 ; 2 à 4 : hommes sains ; 5 : femme conductrice d'une inversion de type 1.

L'analyse par CMA du génome de la patiente 5 a montré la présence d'une délétion de 522 kb (chrX:154,106,184-154,628,722, hg19) englobant 10 gènes et les exons 1 à 22 du gène *F8* (Figure 14 et Tableau 18). Cette délétion est localisée entre les séquences *int22h-1* et *int22h-2*, *int22h-3* n'est pas délété. La MLPA a confirmé la délétion des exons 1 à 22 du gène *F8*.

PCR		Inversion de l'intron 1			CMA		
<i>Int1h-1</i>	<i>Int1h-2</i>	Anomalie	Taille (kb)	Coordonnées génomiques (taille maximale)	Gènes		
Patient 1	1908 pb et 1303 pb	1191 pb	230 52.5	chrX:154,147,182-154,378,045, hg19 chrX: 154,182,749-154,235,426	F8 (exon 1 à intron 14); <i>FUNDC2</i> , <i>CMC4</i> , <i>MTCPI</i> , <i>BRCC3</i> F8 (Intron 1 à exon 11)		
Southern blot		PCR long fragment					
Patient 2	14 kb*, 16 kb, 17.5 kb, 25 kb	10 kb, 11 kb, 11.5 kb, 12 kb	498	chrX:154,224,460-154,722,370, hg19	F8 (exons 1 à 3) <i>FUNDC2</i> , <i>CMC4</i> , <i>MTCPI</i> , <i>BRCC3</i> , <i>VBPI</i> , <i>RAB39B</i> , <i>CLIC2</i> , TMLHE-AS1 copie centromérique, <i>H2AFB3</i> copie intermédiaire, <i>F8AI</i> copie intermédiaire, <i>F8AI</i> copie télomérique, <i>TMLHE-ASI</i> copie télomérique et <i>H2AFB3</i> copie télomérique.		
Patient 3	14 kb*, 16 kb, 17.5 kb	11 kb, 11.5 kb, 12 kb	285 196	chrX: 154,106,243-154,391,634, hg19 chrX: 154,528,419-154,722,429	F8 ( <i>exons 1 à 22</i> ) <i>FUNDC2</i> , <i>CMC4</i> , <i>MTCPI</i> et <i>BRCC3</i> <i>CLIC2</i> (exons 1 et 2), TMLHE-AS1 copie centromérique, <i>H2AFB3</i> copie intermédiaire, <i>F8AI</i> copie intermédiaire, <i>F8AI</i> copie télomérique, <i>TMLHE-ASI</i> copie télomérique et <i>H2AFB3</i> copie télomérique.		
Patient 4	14 kb, 16 kb, 17.5 kb, 15.5 kb, 20 kb	9.6 kb, 11.5 kb, 11 kb, 12 kb, 12.7kb	1302 233	chrX:153,783,581-155,086,461 , hg19 chrX:153,885,800-154,118,678, hg19	<i>CTAG1B</i> , <i>CTAG1A</i> , <i>CTAG2</i> , <i>GAB3</i> , <i>DKCI</i> , <i>MPPI</i> , <i>F8</i> , <i>FUNDC2</i> , <i>CMC4</i> , <i>MTCPI</i> , <i>BRCC3</i> , <i>VBPI</i> , <i>RAB39B</i> , <i>CLIC2</i> , TMLHE-AS1 copie centromérique, <i>H2AFB3</i> copie intermédiaire, <i>F8AI</i> copie intermédiaire, <i>F8AI</i> copie télomérique, <i>TMLHE-ASI</i> copie télomérique et <i>H2AFB3</i> copie télomérique, <i>TMLHE</i> , <i>SRPY-3</i> <i>GAB3</i> , <i>DKCI</i> , <i>MPPI</i> , <i>F8</i> (exons 22 à 26)		
Patient 5	14 kb, 16 kb*, 21.5 kb, 15.5 kb	10 kb, 12 kb, 11 kb, 12.7kb	196 522	chrX: 154,528,419-154,722,429 chrX:154,106,184-154,628,722, hg19	<i>CLIC2</i> (exons 1 et 2), TMLHE-AS1 copie centromérique, <i>H2AFB3</i> copie intermédiaire, <i>F8AI</i> copie intermédiaire, <i>F8AI</i> copie télomérique, <i>TMLHE-ASI</i> copie télomérique et <i>H2AFB3</i> copie télomérique. F8 (exons 1 à 22), <i>FUNDC2</i> , <i>CMC4</i> , <i>MTCPI</i> , <i>BRCC3</i> , <i>VBPI</i> , <i>RAB39B</i> , <i>CLIC2</i> , TMLHE-AS1 copie centromérique, <i>H2AFB3</i> copie intermédiaire, <i>F8AI</i> copie intermédiaire		

**Tableau 18 : Caractérisation de cinq profils anormaux d'inversion 1 et 22.** *Int1h-1* : résultat de l'amplification de la séquence *Int1h-1* ; *Int1h-2* : résultat de l'amplification de la séquence *Int1h-2* ; CMA : Cytogenetic Microarray Analysis. \*la bande présente une intensité environ deux fois supérieure aux autres bandes.

## 5. Discussion

Dans ce travail nous avons décrit et caractérisé par CMA 5 profils anormaux d'inversion de l'intron 1 et de l'intron 22. Dans tous les cas, l'analyse par CMA du génome des patients a montré que des délétions et des duplications de très grande taille étaient associées aux inversions chez ces patients.

La présence d'une duplication était suspectée chez les patients 1, 2 et 4 du fait de la présence de bandes surnuméraires sur le SB (recherche d'une Inv22) ou sur la PCR (recherche d'une Inv1). L'analyse par CMA du génome de ces trois patients a confirmé cette hypothèse et a montré la présence de duplications de très grande taille au niveau de la région Xq28. La taille de ces duplications est comprise entre 230kb (patient 1) à 1302 kb environ (patient 4) et comprend selon les patients soit une partie (patient 1 et 2) soit la totalité (patient 4) du gène *F8* ainsi qu'un nombre variable de gènes voisins du gène *F8* (de 4 autres gènes dupliqués chez le patient 1 à 21 chez le patient 4). De plus, chez les patients 1 et 4, l'analyse par CMA a mis en évidence l'existence de régions tripliquées. Les patients 1, 2 et 4 sont tous des hémophiles A sévères sans inhibiteur et ne présentent aucun signe clinique autre que ceux associés à leur HA.

La duplication de la région située entre les séquences *int22h-1* et *int22h-2* (région contenant 8 gènes : *F8*, *FUNDC2*, *CMC4*, *MTCP1*, *BRCC3*, *VBPI*, *RAB39B* et *CLIC2*) a été décrite par plusieurs études chez des patients non hémophiles mais présentant un déficit mental et des malformations faciales (69,70,170,171). Dans un travail récent, Varmensenille *et al.* ont montré que le gène *RAB39B* était un bon gène candidat pour expliquer le déficit mental observé chez ces patients (171). En effet les patients porteurs de cette duplication présentent des taux de transcrits du gène *RAB39B* très élevés et la surexpression de ce gène chez la souris est responsable de l'apparition d'anomalies du développement du système

nerveux central. Le gène *RAB39B* code une petite GTPase appartenant à une famille de protéines impliquées dans la signalisation intracellulaire intervenant notamment au niveau neuronal. De façon surprenante, tous les patients hémophiles A rapportés à ce jour (à la fois dans ce travail mais aussi dans l'article de Lannoy *et al.* (70)) porteurs à la fois d'une duplication dans la région Xq28 impliquant notamment les gènes *F8*, *RAB39B* et *CLIC2* et d'une Inv22 ou d'une tripllication de certains exons du gène *F8* ne présentent pas de déficit mental ou d'anomalie morphologique. Il est possible que les réarrangements génomiques complexes présents chez ces patients hémophiles A soient responsables d'une diminution du niveau d'expression des gènes dupliqués dont *RAB39B* évitant ainsi l'apparition d'un déficit mental. Malheureusement, nous n'avons pas pu vérifier notre hypothèse car aucun prélèvement d'ARNm n'était disponible pour les patients 1 et 4.

Le patient 3 était un hémophile A sévère avec un taux élevé d'inhibiteur anti-FVIII. Ce patient est décédé des suites d'un accident vasculaire cérébral hémorragique. Le profil anormal en SB de ce patient suggère la présence à la fois d'une Inv22-1 et d'une délétion. L'analyse par CMA de son génome a montré l'existence d'une délétion de 285 kb dans la région Xq28 comprenant les exons 1 à 22 du gène *F8* et 4 autres gènes (*FUNDC2*, *CMC4*, *MTCP1* et *BRCC3*). Miskinyte *et al.* ont montré que la délétion du gène *BRCC3* était une étiologie possible pour le syndrome de Moyamoya lié à l'X (172). Ce syndrome est caractérisé par l'association d'une angiopathie de Moyamoya, d'une petite taille et de malformations faciales. L'angiopathie de Moyamoya est une anomalie de la circulation vasculaire avec une sténose progressive de la partie terminale des artères carotidiennes entraînant l'apparition d'un réseau artériel compensatoire anormal et fragile augmentant le risque d'accident vasculaire cérébral hémorragique ou thrombotique (173). Deux cas d'HA sévères ayant une délétion du gène *BRCC3* ont été rapportés dans la littérature à ce jour : un patient avec un syndrome Moyamoya (174) et un autre patient présentant des signes cliniques

caractéristiques du syndrome Moyamoya lié à l’X mais dont il est impossible d’affirmer le diagnostic du fait de l’absence de données d’imagerie cérébrale (169). Ainsi il est probable que l’accident vasculaire cérébral hémorragique ayant entraîné le décès du patient 3 soit lié à un syndrome Moyamoya non diagnostiqué, d’autant plus que ce patient présente certains signes cliniques évocateurs tels qu’une petite taille et des anomalies morphologiques faciales. De plus, son oncle maternel (hémophile A sévère avec inhibiteurs) est lui aussi décédé d’un accident vasculaire cérébral hémorragique. Enfin, Janczar *et al.* ont montré plus récemment que les femmes conductrices pour les grandes délétions de la région Xq28 comprenant les gènes *F8* et *BRCC3* pouvaient avoir des présentations cliniques variables du syndrome Moyamoya du fait d’un déséquilibre d’inactivation de l’X (175). Une exploration du syndrome Moyamoya a donc été proposée à la sœur et à la mère du patient 3.

Le patient 3 et son oncle maternel ont développé un titre très élevé d’inhibiteurs anti-FVIII. Plusieurs cas de patients hémophiles A sévères ayant à la fois une Inv22 et une délétion du gène *F8* et ayant développé des inhibiteurs ont été rapportés dans la littérature (168,169). Ces observations ne sont pas surprenantes. En effet il est connu que la présence concomitante d’une Inv22 et d’une grande délétion du gène *F8* est associée à un risque beaucoup plus important de développer un inhibiteur que l’Inv22 seule (176).

La patiente 5 est une femme conductrice d’HA à taux bas de FVIII chez laquelle l’analyse par CMA a identifié une délétion de la région comprise entre les séquences *int22h-1* et *int22h-2* (0,5 Mb). De façon intéressante, El-Hattab *et al.* ont décrit la même délétion de 0,5Mb à l’état hétérozygote chez deux femmes ayant des taux normaux de FVIII (69). La patiente 5 est aussi porteuse de l’association de mutation *F8* c.2044G>C (p.Val682Leu) + *F8* c.2887G>A (p.Asp963Asn). Nous avons démontré dans une autre étude que la mutation p.Val682Leu était associée à un phénotype mineur d’hémophilie A ce qui permet d’expliquer le taux de FVIII de la patiente 5 (108). Le mécanisme proposé pour expliquer cette délétion est une

recombinaison homologue non-allélique (NAHR) entre les séquences *int22h-1* et *int22h-2* localisées sur deux chromatides différentes résultant en un crossing-over non équilibré avec formation d'une chromatide portant une duplication et l'autre la délétion réciproque (69,70). La délétion comprend les exons 1 à 22 du gène *F8* et est responsable de l'apparition du fragment de 15,5kb sur le SB correspondant à la séquence hybride *int22h-1/2*. Ainsi, les profils anormaux d'Inv22 observés chez la patiente 5 en SB et LR-PCR ont été faussement interprétés comme conséquence d'une Inv22-2. Comme une des deux femmes décrites par El-Hattab *et al.* a eu deux fausses couches spontanées et qu'à ce jour aucun homme porteur de cette délétion n'a été décrit, El-Hattab *et al.* ont émis l'hypothèse que cette délétion serait létale *in utero* pour les individus masculins. Cette hypothèse suggère donc que la patiente 5, qui est une femme en âge de procréer, pourrait être à risque de fausse couche. Cependant, à ce jour cette patiente a eu une fille et n'a aucun antécédent de fausse couche.

Les patients 2 et 3 sont porteurs de la même duplication de 196 kb située entre l'intron 2 du gène *CLIC2* et *int22h-3* (chrX: 154,528,419-154,722,429). Cette duplication est chevauchante avec la duplication de 180 kb décrite par Lannoy *et al.* chez des hémophiles A sévères porteurs d'Inv22 (64). Cette duplication est responsable de la présence de 5 copies *int22h* au niveau de la région Xq28, l'une dans le gène *F8* (intron 22) et les 4 autres situées dans une région plus télomérique. Cette duplication doit être considérée comme un polymorphisme car elle a été retrouvée à la fois chez des hémophiles A sévères et chez des patients sains (64).

En conclusion, comme plusieurs gènes voisins du gène *F8* sont associés à des pathologies diverses telles que des déficits mentaux et des anomalies cardiovasculaires, tous les patients chez lesquels un réarrangement complexe de la région Xq28 est suspecté devraient être orientés vers un généticien afin d'évaluer l'intérêt d'une étude plus complète de leur génome. De telles investigations doivent être menées dans le cadre d'un conseil génétique pour les



femmes conductrices de ces anomalies afin d'évaluer le risque de transmission de l'hémophilie  
A sévères associée à un « syndrome des gènes contigus ».

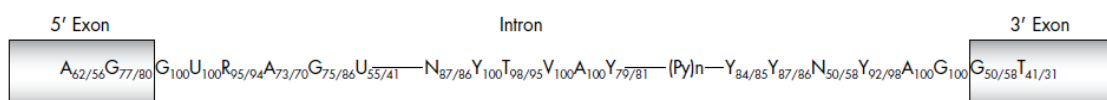
**Section 3. Caractérisation de variants de signification indéterminée localisés  
au niveau des sites d'épissage du gène *F8*.**

## Partie 1 : Rappels bibliographiques sur la réaction d'épissage

Dans les cellules eucaryotes, les ARNm sont synthétisés dans le noyau cellulaire sous forme de pré-ARNm lors de l'étape de transcription. L'excision des parties introniques (régions non-codantes) des pré-ARNm et la jonction des exons (régions codantes) entre eux sont réalisées lors de la réaction d'épissage des pré-ARNm. Cette étape nécessite l'intervention d'un complexe ribonucléique appelé spliceosome. La composition et la structure du spliceosome varient beaucoup au cours de la réaction d'épissage. Il existe dans les cellules eucaryotes deux types de spliceosome : le spliceosome de type U2 catalysant l'excision des introns de type U2 (la majorité des cas) et le spliceosome de type U12 catalysant l'excision des introns de type U12. La totalité des introns du gène *F8* étant de type U2, nous nous intéresserons uniquement à l'épissage des introns par le spliceosome de type U2.

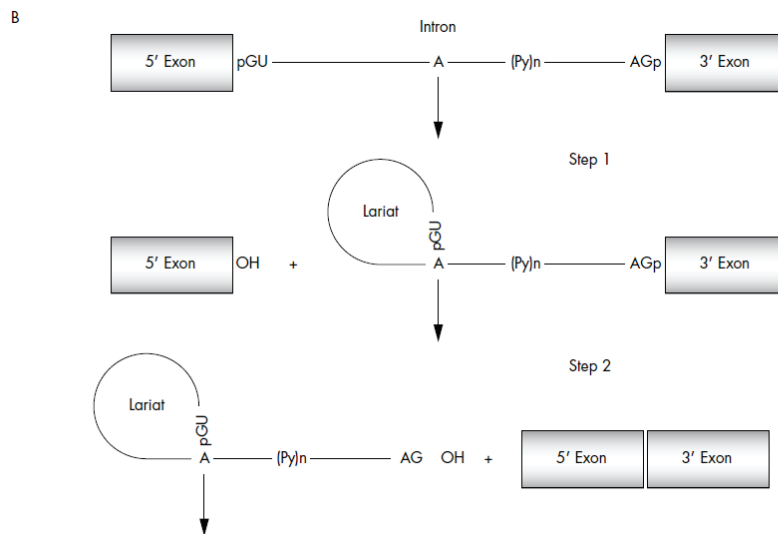
### 1. La réaction d'épissage

L'excision des parties introniques des pré-ARNm et la jonction des exons entre eux sont réalisées lors de la réaction d'épissage des pré-ARNm. Cette réaction consiste en deux réactions successives de trans-estérification aboutissant à l'élimination de l'intron et à la liaison des deux exons successifs (177). Les différents sites nécessaires à la réaction de trans-estérification sont: le site donneur d'épissage (5'ss), le site accepteur d'épissage (3'ss) et le point de branchement (BP). Le site donneur d'épissage est composé du dinucléotide GT invariable en début d'intron. Les nucléotides situés autour de cette séquence consensus ont aussi leur importance mais sont moins conservés. Le site accepteur d'épissage est défini par trois éléments : le BP classiquement localisé 14 à 100 nucléotides en amont de la fin de l'intron, le BP est suivi par un tract polypyrimidines puis par le dinucléotide AG invariable terminant l'intron (Figure 19).



**Figure 19 : Structure des jonctions intron/exon.** Les trois motifs nécessaires à la réaction d'épissage (GU en 5', AG en 3' et le A du point de branchement) sont universellement conservés. Les nucléotides situés autour de ces structures sont importants mais sont plus variables (les chiffres indiquent les fréquences).

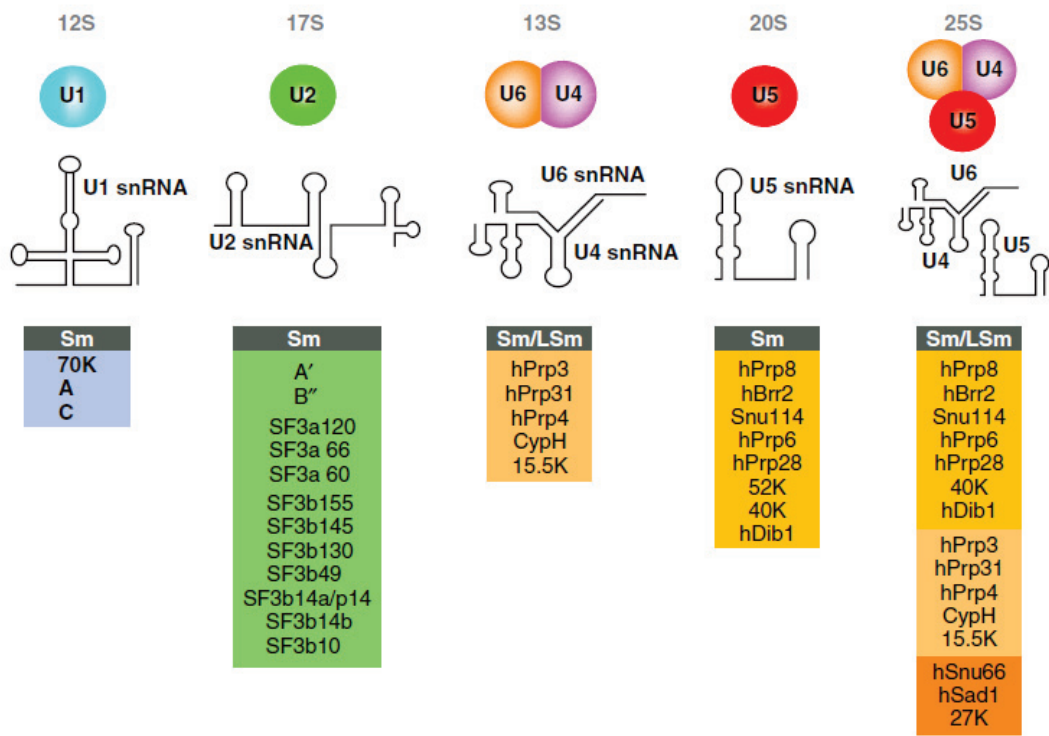
La première étape de la réaction d'épissage consiste en l'attaque par le résidu adénosine du BP de la liaison phosphodiester du site donneur d'épissage. Cette première réaction génère un état intermédiaire avec formation d'une structure en lasso entre l'extrémité 5' de l'intron et le point de branchement et entraîne la création d'un groupement hydroxyle libre au niveau de l'exon localisé en 5' de l'intron. Dans une deuxième étape, le groupe hydroxyle libre de l'exon en 5' de l'intron attaque le site accepteur de l'intron ce qui entraîne la liaison des deux exons entre eux et la libération de l'intron en lasso qui sera rapidement dégradé (Figure 20).



**Figure 20 : La réaction d'épissage consiste en deux réactions de trans-estérification successives (177).** Lors de la première étape, un résidu adénosine (point de branchement) classiquement localisé 14 à 100 nucléotides en amont du site accepteur d'épissage (région 3' terminale de l'intron) va attaquer la liaison phosphodiester au niveau du site donneur d'épissage (extrémité 5' de l'intron). Cette première réaction va générer un état intermédiaire avec formation d'un lasso entre l'extrémité 5' et le point de branchement et la création d'un groupement hydroxyle libre au niveau de l'exon localisé en 5' de l'intron.. Dans une deuxième étape, le groupe hydroxyle libre de l'exon en 5' de l'intron va attaquer le site accepteur de l'intron, entraînant la liaison des deux exons entre eux et la libération de l'intron an lasso qui sera rapidement dégradé. (PY)n : tract de polypyrimidines.

## 2. Le cycle d'assemblage du spliceosome

Le spliceosome de type U2 est composé de 4 petites ribonucleoparticules nucléaires (snRNP) : U1, U2, U5 et U4/U6 et de nombreuses protéines non-snRNP. Chaque snRNP est composée d'un petit ARN nucléaire (snRNA) (deux dans le cas de snRNP U4/U6), d'un ensemble de 7 protéines Sm (B/B', D3, D2, D1, E, F et G) et d'un nombre variable d'autres protéines spécifiques pour chaque snRNP (Figure 21) (178).



**Figure 21 : Composition des principaux snRNPs composant le spliceosome chez l'Homme (178).** Les ensembles de protéines Sm (B/B', D3, D2, D1, E, F et G) et de protéines Lsm (Lsm2-8) sont indiqués par l'indication Sm ou Lsm en haut des cases présentant la liste des protéines formant chacun des snRNP. L'ensemble snRNP U4/U5/U6 est composé de deux ensembles de protéines Sm et d'un ensemble de protéines LSm.

Le recrutement et la libération au cours de la réaction d'épissage de protéines composant le spliceosome forment le cycle d'assemblage du spliceosome (Figure 22).

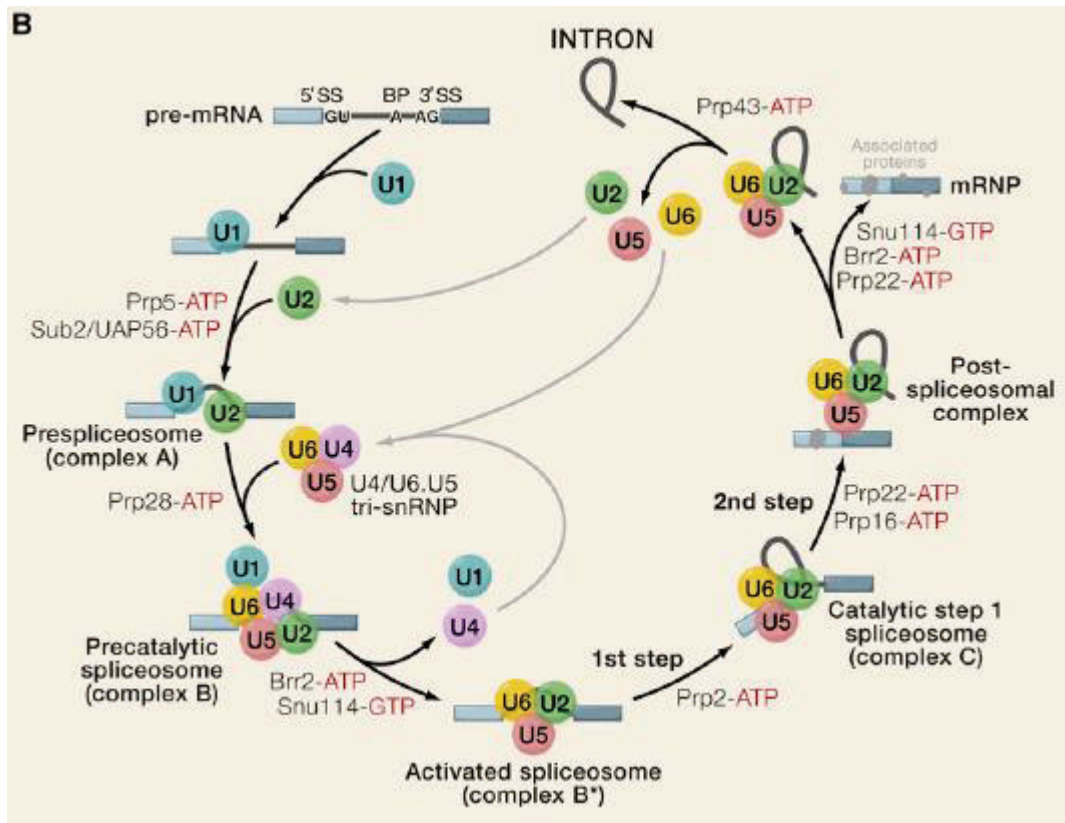
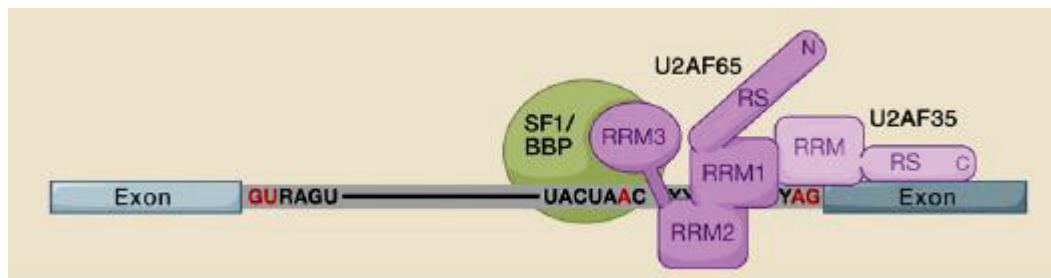


Figure 22 : Cycle d'assemblage du spliceosome (179).

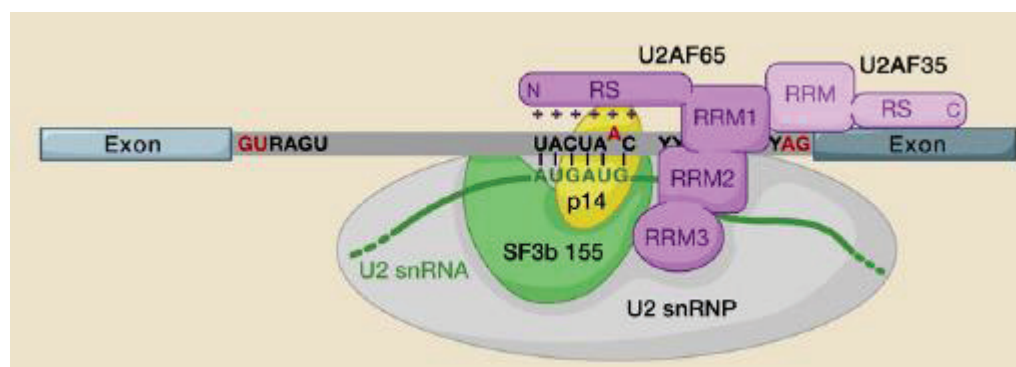
La première étape du cycle du spliceosome consiste en la reconnaissance des sites d'épissage donneur et accepteur par certains composants du spliceosome. Le snRNP U1 est recruté au niveau du site donneur de l'intron (appariement du snRNA U1 au niveau de la séquence du site donneur d'épissage sur le pré-ARN). Comme la majorité des interactions ARN/ARN au sein du spliceosome sont faibles, des protéines sont nécessaires pour stabiliser ces liaisons. Ainsi l'interaction snRNP U1/pré-ARN est stabilisée par des protéines composant le snRNP U1 et des protéines de la famille des protéines riches en sérine et arginine (SR). La phase précoce d'assemblage du spliceosome fait aussi intervenir d'autres protéines non-snRNP. Les protéines SF1/BBP et U2AF (U2 auxiliary factor) se lient respectivement au BP et au tract polypyrimidine. Une coopération entre ces deux protéines est nécessaire pour leur liaison au niveau du site accepteur d'épissage : SF1/BBP interagit avec le domaine RRM3 (RNA recognition motif 3) de la sous-unité de 65kDa de l'hétérodimère U2AF (U2AF65). La sous-unité de 35kDa de U2AF (U2AF35) reconnaît le dinucléotide AG du site accepteur (Figure 23). Les liaisons au pré-ARNm des différents acteurs protéiques décrits ci-dessus

forment le complexe E du spliceosome et jouent un rôle primordial dans la reconnaissance initiale des sites d'épissage par le spliceosome.



**Figure 23 : Interactions moléculaires au niveau du site accepteur d'épissage au sein du complexe E du spliceosome (179).** La protéine SF1/BBP se lie au niveau du site de branchement sur le pré-ARNm (exon en bleu et intron en gris). Les sous-unités de 65kDa et de 35kDa du facteur auxillaire U2 (U2AF) lient respectivement le tract polypyrimidine et le dinucléotide AG. La sous-unité U2AF65 interagit à la fois avec SF1/BBP et U2AF35.

Après la formation du complexe E, le snRNA U2 s'apparie au niveau du site de branchement du pré-ARNm de manière ATP-dépendante entraînant la formation du complexe A. Cette interaction ARN/ARN est stabilisée à la fois par une protéine hétérodimérique appartenant au snRNP U2 : SF3a et SF3b et par la protéine U2AF65. L'appariement du snRNP U2 au niveau du site de branchement entraîne la libération de SF1/BBP. Les fonctions de SF1/BBP sont assurées au sein du complexe A par l'association SF3b14a/p14 pour l'interaction avec le BP et par SF3b155 au niveau de l'interaction avec U2AF65 (Figure 24).



**Figure 24 : Interactions moléculaires au niveau du site accepteur d'épissage au sein du complexe A du spliceosome (179).** La protéine SF1/BBP est libérée et remplacée dans ces fonctions par la protéine du snRNP U2 p14 pour l'interaction avec le point de branchement et par la protéine SF3b155 pour l'interaction avec U2AF65.

A la suite du complexe A, les snRNP U4/U6 et U5, préalablement rassemblés en un complexe U4/U6.U5 tri-snRNP, sont recrutés au sein du splicéosome pour former le complexe B. Bien que tous les snRNP soient présents au sein du complexe B, ce dernier est inactif. Des réarrangements

conformationnels et des modifications de la composition du splicéosome sont nécessaires à l'acquisition d'une activité catalytique pour les réactions de trans-estérification. Durant l'activation du splicéosome (complexe B'), les snRNP U1 et U4 sont libérés. Le complexe B' catalyse la première réaction de trans-estérification entraînant la formation du complexe C. Après de nouveaux réarrangements au sein du splicéosome, la seconde trans-estérification a lieu. A la suite de cette étape, l'ARNm est libéré et les snRNP U2, U5 et U6 sont recyclés pour un nouveau cycle.

Les deux premières étapes du cycle du splicéosome (formation des complexe E et A) illustrent plusieurs principes qui caractérisent la réaction d'épissage. Premièrement, la liaison de façon séquentielle au niveau des différents sites réactionnels des introns (site accepteur, donneur, points de branchement et tract polypyrimidine) de snRNA et de facteurs protéiques assurent la correcte reconnaissance des sites d'épissage. Deuxièmement, au sein du splicéosome, les interactions moléculaires essentielles à la réaction d'épissage sont des interactions faibles, ces dernières étant renforcées par des interactions avec de multiples facteurs auxiliaires. Ce phénomène est indispensable à la flexibilité et à la régulation du splicéosome.

### **3. Eléments participant à la reconnaissance de l'exon par le splicéosome**

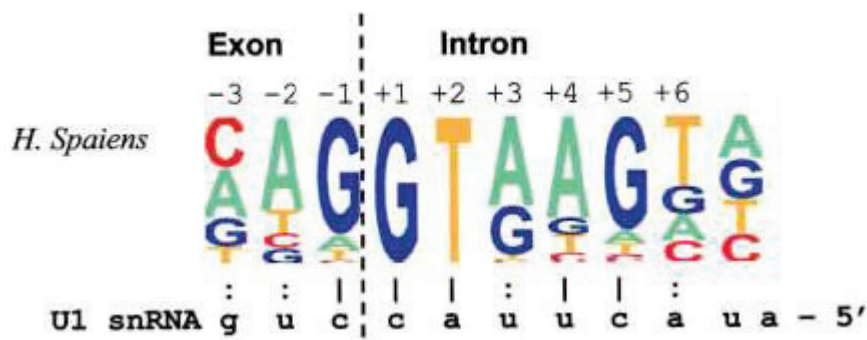
Parmi les 25 000 gènes présents dans le génome humain, des épissages alternatifs ont été mis en évidence pour plus de 90% d'entre eux (180). L'existence d'épissage alternatif pour un même pré-ARNm permet la production de multiple isoformes d'une même protéine. Ce phénomène a permis une diversification très importante du protéome chez les organismes eucaryotes évolués à partir d'un nombre limité de gènes (181). Parce que la régulation de l'épissage alternatif des pré-ARNm permet de déterminer à quel moment et dans quel type cellulaire est produite une certaine isoforme protéique, tout changement dans ce processus module de nombreuses activités cellulaires différentes. Les systèmes de régulation de l'épissage sont donc un élément clé du métabolisme cellulaire. De nombreux éléments participent aux systèmes de régulation de ce processus.



## 1. Les sites d'épissage

### a) Le site donneur d'épissage (5'ss)

Le motif consensus du site donneur d'épissage consiste en 9 nucléotides, CAGGTAAGT, localisés à cheval sur la jonction exon/intron. Les trois premiers (numérotés de -3 à -1) sont exoniques alors que les 6 autres sont introniques (numérotés +1 à +6). La séquence consensus du site donneur d'épissage est parfaitement complémentaire de la séquence du snRNA U1 (Figure 25). Ainsi, la force d'un site donneur d'épissage est dépendante de la complémentarité de sa séquence avec le snRNA U1. La conservation des nucléotides de la séquence consensus du site donneur permet de déterminer l'importance de chaque nucléotide dans la liaison avec le snRNA U1. Le dinucléotide GT en position +1 et +2 est constamment présent au niveau des sites donneurs des introns de type U2 et est indispensable pour la liaison du snRNA U1 (182). Les nucléotides G en position -1 et +5 sont aussi très conservés (environ dans 80 % des cas) car ils permettent la formation de trois liaisons hydrogènes chacun avec le snRNA U1 (183). Au niveau de la position +3, un A ou G sont présents dans 80% des cas également car ces deux nucléotides peuvent s'apparier à la pseudo-uridine du snRNA U1. Les autres positions du site donneur sont plus dégénérées (184).



**Figure 25 : La séquence consensus du site donneur d'épissage (183).** La taille des lettres représente la fréquence du nucléotide en question à une position donnée.

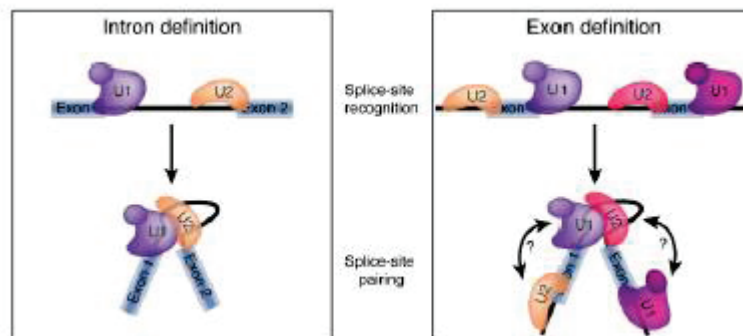
D'autres snRNA reconnaissent aussi le site donneur d'épissage. Le snRNA U6 se lie au niveau des positions +2, +5 et +6 et le snRNA U5 se fixe au niveau des positions -1 et -2 (185).

## b) Le site accepteur d'épissage (3'ss)

A la différence du site 5'ss, le site accepteur est défini par plusieurs éléments contenus classiquement dans les 40 derniers nucléotides de l'intron. Le premier de ces éléments est la jonction intron/exon dont la séquence consensus est YAG/G (Y= C ou T). Le dinucléotide AG (positions -1 et -2 du site accepteur) est invariable et indispensable à la fixation de la sous-unité U2AF35 du snRNP U2 (186). Le deuxième élément est le tract polypyrimidine (région riche en C ou T). Le dernier élément du 3'ss est le BP. La séquence consensus du BP est très dégénérée chez les eucaryotes supérieurs (YTNAV ; Y = C ou T ; N = A, C, G ou T) (187). Seul le A est très conservé car il intervient dans la première réaction de trans-estérification. Le BP est généralement localisé entre 14 et 40 nucléotides en amont du site accepteur d'épissage. Il peut parfois se trouver jusqu'à 100 nucléotides. Le snRNA U2 se fixe par complémentarité au niveau du BP lors de l'assemblage du splicéosome. Le splicéosome sélectionne le premier AG disponible en aval du point de branchement et du tract polypyrimidine pour la réaction d'épissage. Du fait de l'encombrement stérique lié aux complexes protéiques fixés au niveau du BP, le dinucléotide AG doit être néanmoins situé au minimum à 19 pb du point de branchement (188).

## 2. Rôle de l'architecture intron / exon

L'architecture exon/intron a une influence sur la reconnaissance des sites d'épissage (189). La reconnaissance des sites d'épissage par le splicéosome est facilitée si les introns ou les exons sont de petite taille. Deux modèles de reconnaissance des sites d'épissage par le splicéosome ont été proposés (190). Pour les introns de dépassant pas 250 nucléotides, le splicéosome reconnaît les sites d'épissage selon le « modèle de l'intron ». Dans ce cas le splicéosome s'assemble au niveau des sites d'épissage situés de part et d'autre de l'intron. Pour les autres cas (introns de plus de 250 pb), c'est le modèle de l'exon qui prime : le splicéosome s'assemble au niveau des sites d'épissage situés de part et d'autre de l'exon (Figure 26).



**Figure 26 : Schématisation des deux modèles de reconnaissance des sites d'épissage par le spliceosome (190).**

Le « modèle de l'intron » est plus efficace pour la reconnaissance des sites d'épissage par le spliceosome. Ainsi de nombreux exons du génome humain flanqués par des introns courts sont correctement épissés en dépit de site d'épissage très faible (séquence éloignée de la séquence consensus). *A contrario*, les exons flanqués par des introns de grande taille sont souvent sautés lors d'épissage alternatif. Enfin des données expérimentales semblent indiquer que la taille de l'intron en amont de l'exon est plus importante pour définir un épissage alternatif que celle de l'intron en aval (191).

### 3. La structure secondaire des pré-ARNm

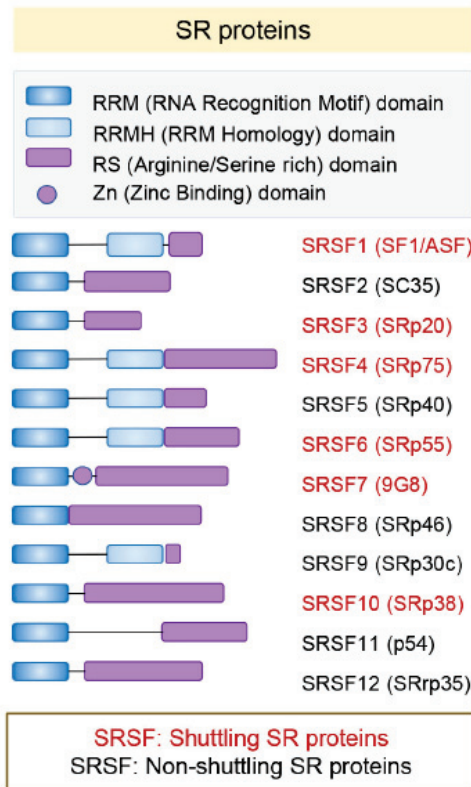
Les pré-ARNm peuvent adopter des structures secondaires les maintenant sous forme double brin localement (formation de boucles ou d'épingles à cheveux). En fonction de leur stabilité, ces structures peuvent persister dans le temps et interférer ou moduler la réaction d'épissage soit en activant soit en inhibant l'assemblage du spliceosome. En effet la reconnaissance des sites d'épissage et les fonctions activatrices ou inhibitrices de l'épissage de nombreuses protéines de régulation sont dépendantes d'interactions protéine/ARNm simple brin. Ainsi les structures locales des ARNm peuvent soit inhiber l'assemblage du spliceosome en masquant un site d'épissage ou des séquences régulatrices activatrices ou au contraire favoriser son assemblage en masquant une séquence inhibitrice (190). Les structures secondaires d'épissage peuvent aussi faire varier la distance relative entre différents signaux consensus d'épissage sur le pré-ARNm afin de moduler leur reconnaissance par le spliceosome (192).

#### 4. Les séquences cis-régulatrices de l'épissage

Les introns du génome humain contiennent de nombreuses régions dont la séquence se rapproche des séquences consensus des sites d'épissage. En dépit de la présence de ces séquences leurres, le splicéosome ne reconnaît que les sites d'épissage authentiques. Ceci est lié à la présence de nombreuses séquences régulatrices de l'épissage. Ces séquences sont appelées séquences activatrices ou inhibitrices exoniques (ESE pour exonic splicing enhancers et ESS pour exonic splicing silencers) ou introniques (ISE pour intronic splicing enhancers et ISS pour intronic splicing silencers).

##### a) Les séquences exoniques activatrices (ESE)

La majorité des exons du génome humain contiennent des séquences régulatrices de type ESE. La plupart de ces séquences fonctionnent en recrutant des éléments trans-régulateurs appartenant à la famille des protéines SR (protéines riches en sérines/arginines). Ces protéines sont constituées d'un domaine riche en sérines et en arginines à leur extrémité C-terminale (domaine SR) et d'un domaine de reconnaissance de l'ARN à leur extrémité N-terminale (domaine RRM pour RNA Recognition Motif) (Figure 27) (193). Typiquement ces protéines se fixent sur le pré-ARNm au niveau d'une séquence ESE via leur domaine RRM et interagissent avec d'autres protéines par l'intermédiaire de leur domaine RS afin de faciliter l'assemblage du splicéosome. Les séquences consensus des ESE reconnues par les domaines RRM sont très dégénérées et chaque protéine SR peut se lier à différents motifs (Tableau 19) (194).



**Figure 27 : Liste et structure des protéines SR (193).** Les protéines indiquées en noires sont localisées exclusivement dans le noyau cellulaire et celles en rouges sont localisées dans le noyau et dans le cytoplasme.

Protein	Binding site <sup>a</sup>	Method <sup>b</sup>	Reference
SF2/ASF	RGAAGAAC AGGACRRAGC	SELEX	Tacke & Manley, 1995
SC35	SRSASGA AGSAGAGUA GUUCGAGUA UGUUCSAGWU GWUWCCUGCUA GGGUAUGCUG GAGCAGUAGKS AGGAGAU GRYYCSYR	Functional SELEX	Liu et al., 1998 Tacke & Manley, 1995
9G8	AGACKACGAY ACGAGAGAY	SELEX	Cavaloc et al., 1999
SRp40	UGGGAGCRGUYRGCUCGY ACDGS	SELEX Functional	Tacke et al., 1997 Liu et al., 1998
SRp55	USCGKM	SELEX	Liu et al., 1998
B52	GRUCAACCDNNGCGAACNG	SELEX	Shi et al., 1997
hTra2 $\beta$	(GAA) <sub>n</sub>	SELEX	Tacke et al., 1998

<sup>a</sup>N: any nucleotide; R: purine; Y: pyrimidine; S: G or C; K: U or G; W: A or U; D: A, G, or U; M: A or C.

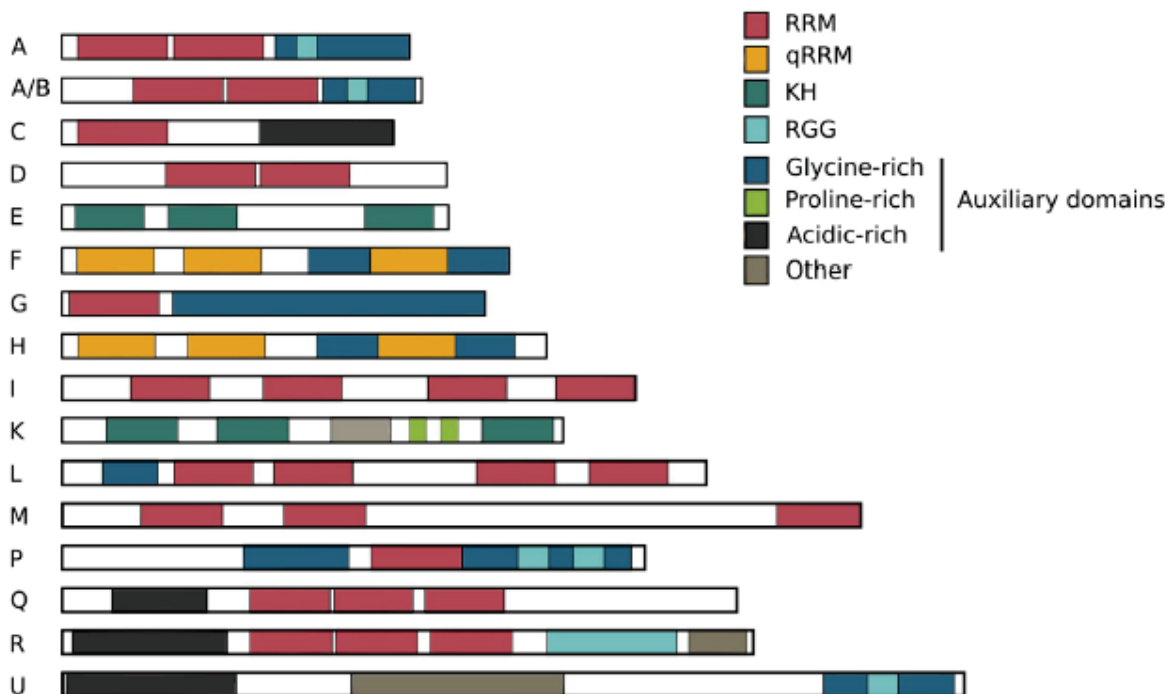
<sup>b</sup>"SELEX" indicates that the RNA sequence was determined to be a high affinity binding site for a purified SR protein; "functional" indicates that the RNA sequence was determined to function as an SR protein-specific splicing enhancer.

**Tableau 19 : Séquences identifiées comme étant des sites de liaison des protéines SR (194)**

b) Les séquences exoniques inhibitrices (ESS)

Les séquences ESS sont reconnues par des protéines ayant des fonctions inhibitrices sur l'épissage. Ces protéines appartiennent à la famille hnRNP (heterogeneous nuclear ribonucleoproteins). Cette famille regroupe environ 30 protéines (hnRNP A1 à hnRNP U), certaines pouvant se lier à l'ARNm (hnRNP majeures) et d'autres non (hnRNP mineures). Les hnRNP mineures sont moins exprimées et ont pour fonction la régulation des hnRNP majeures (195).

Trois types de domaine de liaison à l'ARNm ont été identifiés chez les hnRNP (Figure 28) : les domaines de type RRM, les domaines riches en glycine (RGG) et les domaines de type KH (K-homology, domaine identifié en premier sur l'hnRNP K) (196). Les hnRNP possèdent aussi des domaines auxiliaires tels que les domaines riches en prolines, glycines ou en acides aminés acides. Ces domaines permettent les interactions protéines/protéines avec d'autres hnRNP ou d'autres types de protéines. La grande variété de structure des différentes hnRNP est responsable d'une grande diversité dans les fonctions de ces protéines (196).



**Figure 28 : Les protéines hnRNP (196).** Les différentes protéines de cette famille ont un poids moléculaire compris entre 34 et 120 kDa. De nombreuses hnRNP ont été retrouvées au sein des mêmes complexes protéiques ce qui suggère que plusieurs hnRNP pourraient avoir la même fonction.

c) Les séquences introniques activatrices (ISE) et inhibitrices (ISS)

Un très grand nombre de séquences régulatrices introniques ISE ou ISS ont été décrites à ce jour (197). L'une des ISE la mieux connue est le triplet GGG ou les tracts de G ( $G_n$ ,  $n>3$ ), souvent présents sous la forme de motifs répétés, et qui augmentent la reconnaissance par le spliceosome des sites d'épissage situés à proximité (198). Ce type d'ISE est commun dans les introns riches en GC (199). Les répétitions introniques de CA ont été décrites comme des ISE dans plusieurs cas. Ces séquences favorisent l'épissage de l'exon situé en amont probablement via la fixation de l'hnRNP L (200,201). D'autres séquences décrites comme des ISE ont été rapportées telles que l'héxanucléotide UGCAUG et le motif YCAY (Y= C ou T) (202,203). Il est important de noter que certaines séquences (comme par exemple les répétitions de CA ou le motif YCAY), en fonction de leur localisation, peuvent fonctionner comme des ISE ou des ISS (200). Enfin plusieurs séquences ISS sont connues comme par exemple le motif UCUU qui est reconnu par l'hnRNP I (aussi connu sous le nom de PTB pour Polypyrimidine tract Binding). Cette protéine se comporte comme un répresseur de l'épissage notamment en empêchant les composants du spliceosome de se fixer sur le tract polypyrimidine (204).

## Partie 2. Etude *in vitro* de l'impact sur l'épissage de 26 variations du gène *F8*

### 1. Résumé

**Contexte :** La pathogénicité des variations nucléotidiques localisées au niveau des sites d'épissage est difficile à évaluer. A ce jour, seul quelques patients hémophiles A porteurs de telles variations ont pu bénéficier d'une analyse des transcrits du gène *F8* afin de déterminer l'impact de ces variations sur l'épissage du gène.

**Objectifs :** Etudier *in vitro* l'impact de la totalité des variations nucléotidiques potentiellement responsables d'anomalies d'épissage trouvées dans la cohorte lyonnaise de familles d'hémophiles A.

**Matériels et méthodes :** L'impact sur l'épissage de 26 variations nucléotidiques a été déterminé à l'aide d'algorithmes *in silico* (MaxEntScan et Human splicing Finder) et par technique minigène (réalisée dans deux modèles cellulaires Huh7 et HeLa).

**Résultats :** Une anomalie de l'épissage du gène *F8* a été démontrée pour 21 des 26 variants étudiés. Ces mutations sont responsables le plus souvent d'un saut d'exon (n=13/19). Les autres mécanismes moléculaires trouvés sont l'activation de sites cryptiques d'épissage (n=5), l'utilisation de site *de novo* d'épissage (n=2) ou l'insertion totale de l'intron (n=1). Dans cette étude, 5 mutations silencieuses du gène *F8* ont été reclassifiées en mutations d'épissage : c.1167A>G (p.Gln389Gln), c.1569G>T (p.Leu523Leu), c.1752G>A (p.Gln584Gln), c.5586G>A (p.Leu1862Leu), and c.6066C>T (p.Gly2022Gly). La pathogénicité de 5 variants étudiés reste par contre encore incertaine (c.222G>A, c.237C>T, c.240C>T, c.2113+5\_2113+8del et c.2113+5G>A).

**Discussion :** La démonstration de la pathogénicité d'un variant d'épissage permettant d'améliorer le conseil génétique donné à la famille. Ce type d'étude devrait être réalisé pour tous les patients porteurs de variations affectant potentiellement l'épissage du gène *F8* pour lesquels aucun échantillon d'ARNm n'est disponible.



## 2. Introduction

L'hémophilie A (HA) est causée par une grande variété d'anomalies génétiques et plus de 2000 mutations ont été reportées dans les bases de données internationales. (FVIII variant database, <http://www.factorviii-db.org/>; CHAMP Mutation List Database, [www.cdc.gov/ncbddd/hemophilia/champs.html](http://www.cdc.gov/ncbddd/hemophilia/champs.html); Mai 2017). Approximativement 8% des variants répertoriés sont classés comme affectant potentiellement l'épissage du gène *F8*. Cependant, en absence d'étude de l'effet de ces variations sur l'épissage, la pathogénicité de ces anomalies est difficile à évaluer.

Les mutations d'épissage peuvent affecter l'épissage d'un gène soit en diminuant la force des sites d'épissage physiologiques donneurs ou accepteurs, soit en créant un nouveau site d'épissage, ou alors en modifiant des séquences régulatrices. Il est classiquement admis que les mutations affectant les séquences invariables GT/AG flanquant les exons sont responsables d'anomalies d'épissage. A l'inverse, l'impact sur l'épissage des variations génétiques touchant des positions plus éloignées et plus variables des sites d'épissages est difficilement prévisible. Pour aider les généticiens moléculaires, des algorithmes informatiques d'aide à l'interprétation des variants d'épissage ont été développés. Ces algorithmes donnent un score au site d'épissage physiologique (force du site) et le comparent à celui du site d'épissage muté. Cependant l'interprétation des résultats donnés par ces logiciels doit être faite avec prudence : aucun logiciel n'a une sensibilité et une spécificité de 100% et ils ne donnent aucune indication sur le type d'altération entraînée par la mutation (effet total ou partiel ; saut d'exon ou activation d'un site cryptique d'épissage) (205). Idéalement, la démonstration de la pathogénicité d'un variant d'épissage se fait par l'association de l'étude *in silico* et de l'analyse de l'ARNm du patient par RT-PCR et séquençage. Dans le cas de l'hémophilie A, le FVIII étant essentiellement produit par le foie (206,207), il est impossible d'obtenir des prélèvements d'ARNm issus du tissu producteur de FVIII. A ce jour, les études de l'impact de mutations sur l'épissage du gène *F8* ont toutes été réalisées sur les ARNm ectopiques leucocytaires (124,125,205,208–220). Cependant, l'étude des transcrits du gène *F8* nécessite de réaliser le prélèvement sanguin en présence

de conservateurs des ARNm (prélèvement sur tube Paxgene). Ces tubes étant peu disponibles au niveau des services de consultation, peu de patients bénéficient de ce type d'étude.

L'objectif de ce travail était de statuer sur la pathogénicité de 26 variations génétiques potentiellement responsables d'anomalie d'épissage. Ces anomalies ont été étudiées à la fois à l'aide de méthodes *in silico* et *in vitro* par technique minigène.

### **3. Matériels et méthodes**

#### **1. Sélection des variations nucléotidiques à analyser par étude fonctionnelle**

Tous les patients inclus dans ce travail ont signé un consentement éclairé pour la réalisation d'analyses génétiques. Parmi notre cohorte de 2074 familles chez lesquelles des analyses génétiques du gène *F8* ont été réalisées dans le cadre du diagnostic d'HA, 44 mutations pouvant potentiellement avoir un impact sur l'épissage ont été identifiées chez 52 familles différentes (soit 2,5 % des familles de notre cohorte). Ces 44 variants consistaient en 8 substitutions silencieuses et 36 variations génétiques situées au niveau des sites d'épissage. Dix-huit de ces substitutions étaient localisées au niveau des séquences invariantes AG/GT. Ces mutations ont été exclues de cette études car elles ont été considérées comme certainement responsables du phénotype clinique des patients (ces mutations ont été mises en évidence uniquement chez des hémophiles A sévères). Une étude fonctionnelle de l'impact sur l'épissage a donc été réalisée sur les 26 variants restants. Les caractéristiques des variations génétiques étudiées dans ce travail sont présentées dans le Tableau 20.

Patient	Sexe	FVIII:C (%)	Phénotype	Mutation <sup>a</sup>	Exon/Intron	Position exonique	Protéine <sup>b</sup>
Variations nucléotidiques au niveau du site donneur d'épissage							
1	M	17	Mineur	c.143+3_143+4insG	1	-	-
2	M	44	Mineur	c.143+4A>G	1	-	-
3	M	30	Mineur	c.222G>A*	2	78/122	p.Thr74Thr
4	M	36	Mineur	c.237C>T	2	94/122	p.Asn79Asn
5	M	29	Mineur	c.240C>T	2	97/122	p.Ile80Ile
6	M	30	Mineur	c.388+5G>A*	3	-	-
7	M	40	Mineur	c.601+5G>A*	4	-	-
8	M	<1	Sévère	c.[670G>A;3674A>G]	5	69/69	p.[Gly224Arg;Gln1225Arg]
9	M	15	Mineur	c.[787+5G>A;787+12_787+13del]	6	-	-
10	M	28	Mineur	c.1009G>A	7	222/222	p.Asp337Asn
11	M	7	Mineur	c.[-51G>A;1752G>A]	11	215/215	p.Gln584Gln
12	M	24	Mineur	c.2113+5_2113+8del	13	-	-
13	M	39	Mineur	c.2113+5G>A	13	-	-
14	M	36	Mineur	c.5373+5G>T	15	-	-
15	M	14	Mineur	c.5586G>A	16	213/213	p.Leu1862Leu
16	M	24	Mineur	c.5998+3A>T	18	-	-
17	M	24	Mineur	c.5998+3A>C	18	-	-
18	M	30	Mineur	c.6066C>T	19	251/300	p.Gly2003Gly
Variations nucléotidiques au niveau du site accepteur d'épissage							
19	M	17	Mineur	c.144-3C>T	1	-	-
20	M	22	Mineur	c.144-11T>G	1	-	-
21	M	3	Modéré	c.144A>T	2	1/122	p.Arg48Ser
22	M	25	Mineur	c.671-11T>C	5	-	-
23	M	27	Mineur	c.1167A>G	8	158/262	p.Gln389Gln
24	M	3	Modéré	c.1538-18G>A*	10	-	-
25	M	4	Modéré	c.1538-3T>G	10	-	-
26	M	5	Mineur	c.1569G>T*	11	32/215	p.Leu523Leu
27	M	8	Mineur				

**Tableau 20 : Variations nucléotidique analysées *in vitro* dans cette étude.** <sup>a</sup> Quand deux variations nucléotidiques ont été trouvées chez un même patient, le variant étudié *in vitro* est indiqué en gras. <sup>b</sup> L'impact du variant au niveau protéique est

donné uniquement pour les variants exoniques. Quand deux variations nucléotidiques ont été trouvées chez un même patient, le variant étudié *in vitro* est indiqué en gras. \* variants pour lesquels une étude des ARNm ectopiques de patients a été publiée.

## 2. Analyses *in silico*

Les études *in silico* de prédiction d'impact sur l'épissage ont été réalisées à l'aide de deux logiciels basés sur des principes différents (MaxEntScan et Human Splicing Finder [HSF]) et disponibles dans Alamut 2.4.0. Les paramètres par défaut ont été utilisés pour chacun des deux algorithmes.

HSF utilise le modèle de la matrice de poids (Weight Matrice Model) (221). Ce concept permet d'établir la force intrinsèque d'un site d'épissage en comparant les similitudes entre la séquence d'intérêt et la séquence consensus du site correspondant (donneur ou accepteur). Ce modèle prend en compte l'importance relative de chaque nucléotide à chaque position. A partir de cette matrice, un score est calculé pour le site d'épissage. Ce score est compris entre 0 et 100, où 100 représente une similarité parfaite avec la séquence consensus. Dans ce modèle, le site donneur d'épissage s'étend des positions -3 à +6 et le site accepteur s'étend des positions -14 à +1.

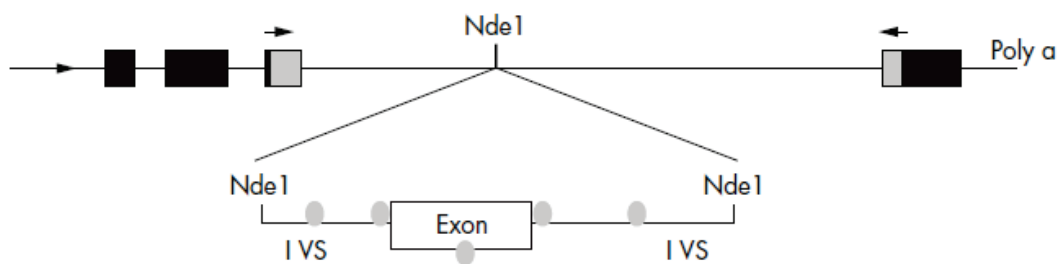
MaxEntScan utilise le modèle du maximum d'entropie (222). Ce modèle prend en compte à la fois l'importance relative de chaque nucléotide à chaque position mais aussi les nucléotides adjacents et non adjacents. Dans ce modèle, le site donneur d'épissage s'étend des positions -3 à +6 et le site accepteur s'étend des positions -20 à +3.

En l'absence de recommandations pour l'interprétation de ces algorithmes, les règles suivantes ont été définies : un variant a été considéré comme affectant l'épissage si les prédictions indiquaient une différences d'au moins 15% pour MaxEntScan et d'au moins 5% pour HSF (223) entre la force du site physiologique et celle su site muté. Dans le cas des sites *de novo* d'épissage, le variant a été considéré comme délétère si la force du site cryptique était égale ou supérieure à celle du site naturel d'épissage.

### 3. Analyse *in vitro* de l'épissage.

#### a) Description du plasmide utilisé

Le minigène (pTB2) utilisé dans ce travail a été décrit par Baralle *et al* en 2005 (177). Ce minigène est un plasmide contenant une version simplifiée d'un gène (Figure 29). Il est composé à son extrémité 5' du promoteur du gène de l' $\alpha$ -globine ainsi que d'une séquence régulatrice « enhancer » de type SV40 permettant la transcription du minigène par la polymérase II dans les cellules transfectées. Cette région est suivie d'une série d'exons et d'introns issus soit du gène de l' $\alpha$ -globine soit du gène de la fibronectine. L'extrémité 3' est constituée du signal de polyadénylation du gène de l' $\alpha$ -globine. Un exon d'intérêt peut être inséré dans ce minigène au niveau du site de restriction unique *NdeI* situé dans l'intron 3 du minigène.



**Figure 29 :** Représentation schématique du minigène utilisé dans ce travail (177). La région 5' est constituée du promoteur de l' $\alpha$ -globine et d'une séquence SV40 (indiquée par la flèche en début de construction). Cette région est suivie par une succession d'exons (rectangles) et d'introns (lignes). Les exons sont issus soit du gène de l' $\alpha$ -globine (rectangles noirs) soit du gène de la fibronectine (rectangles gris). Le signal de polyadénylation utilisé appartient au gène de l' $\alpha$ -globine. L'exon d'intérêt ainsi que les régions introniques flanquantes sont insérées au niveau du site de restriction unique *NdeI*. Les flèches situées de part et d'autre de ce site de restriction correspondent aux amorces de PCR  $\alpha$ 23G et  $\beta$ 2F utilisées pour l'étude de l'épissage du minigène.

Pour l'étude des mutations touchant le site donneur du 1<sup>er</sup> exon du gène *F8*, une version modifiée de ce vecteur a été utilisée. Les 214 paires de bases du premier exon et les 30 paires de bases du premier intron du minigène ont été éliminées afin de permettre l'insertion de l'exon d'intérêt au niveau du site de restriction unique *SmaI* (224).

Le vecteur pTB2 contient aussi un gène de résistance à l'ampicilline.

b) Construction des vecteurs minigènes pour chaque exon du gène *F8* étudié.

Dans un premier temps, les exons d'intérêt et environ 200 pb introniques de part et d'autre de ces exons ont été amplifiés à partir de l'ADN génomique des patients et des témoins. Pour cette étape, une polymérase de haute-fidélité a été utilisée (Optimase polymerase, Transgenomic, Omaha, NE, Etats-Unis). Les amorces utilisées contenaient à leur extrémité 5' une séquence spécifique permettant l'insertion des produits de PCR au niveau du site *NdeI* :

- Queue de l'amorce sens : 5'-CTAACAGCCACATATG-amorce-3'
- Queue de l'amorce antisens : 5'-CCCCCTCGACCATATG-amorce-3'

Dans le cas de l'exon 1 du gène *F8*, les queues spécifiques utilisées pour l'insertion au niveau du site *SmaI* étaient les suivantes :

- Queue de l'amorce sens : 5'-TCGAGGATCTCCCCGGG-amorce-3'
- Queue de l'amorce antisens : 5'-GCGGGCGAGGAGCCCCGGG-amorce-3'

La liste des amorces utilisées pour la construction des différents minigènes est présentée en annexe de cette section. La composition du mélange réactionnel et les conditions de PCR utilisées sont présentées dans les Tableau 21 et Tableau 22 :

Réactifs	Volume (µl)
Tampon Optimase 10X	5
ADN patient ou témoin	100 ng
Amorces sens [10pM]	1,25
Amorces antisens [10pM]	1.25
dNTP [2,5 mM de chaque dNTP]	2
Optimase Polymerase [2,5U/µl]	0,5
Eau	Qsp 50

**Tableau 21 : Composition du mélange réactionnel pour la réaction de PCR optimase.**

	Température (°C)	Temps (sec)	Nombre de cycles
Dénaturation	94	20	14
Hybridation	57	40	
Elongation	72	45	
Dénaturation	94	20	25
Hybridation	50	40	
Elongation	72	45	

**Tableau 22 : Conditions de PCR pour la réaction de PCR optimase.**

Après la réaction de PCR, la taille du produit amplifié et sa pureté ont été vérifiées par migration sur gel d'agarose 1%. Les produits de PCR ont ensuite été purifiés sur colonne à l'aide du kit NucleoSpin Extract II (Clontech laboratories, Mountain View, CA, Etats-Unis), repris dans 30 µl de tampon TE et dosés.

Parallèlement, 1µg de plasmide pTB2 a été digéré par l'enzyme de restriction *NdeI* (ou *SmaI* dans le cas de l'exon 1) une nuit à 37°C puis l'enzyme a été inactivée par chauffage 15 minutes à 65°C. La digestion du plasmide a été contrôlée par migration sur gel d'agarose 0,8% (1h30, 80V, le plasmide PTB2 digéré fait 3454 pb).

Pour l'insertion de l'exon d'intérêt dans le plasmide, le kit In-fusion HD PCR cloning Kit with cloning enhancer II (Clontech) a été utilisé en suivant les recommandations du fournisseur. Brièvement, dans une première étape 250 ng de produit de PCR ont été pré-incubés avec 2µl de cloning enhancer (Qsp 5µl) 20 minutes à 37°C puis 15 minutes à 80°C. Ensuite, à 100 ng de produits de PCR pré-incubés ont été ajoutés 200 ng de vecteur digéré, 2 µl de 5X In-Fusion HD Enzyme Premix et de l'eau (Qsp 10µl). Le mélange réactionnel a été incubé 15 minutes à 37°C puis 15 minutes à 50°C. A la fin, 40 µl de tampon TE ont été ajoutés.

### c) Production, purification et vérification des vecteurs minigènes

Cinquante microlitres de suspension de bactéries DH5α (DO comprise entre 0,2 et 0,4 à 600nm) préalablement rendues compétentes ont été transformées par choc thermique (30 minutes dans la glace, 45 secondes au bain-marie à 42°C puis 2 minutes dans la glace) avec 2 µl de produit de réaction In-fusion. Après ajout de 250 µl de milieu S.O.C (Life technologies, Saint-Aubin, France) préchauffé à 37°C, les tubes ont été incubés à 37°C sous agitation pendant 1 heure. Ensuite, 100 µl de cette suspension bactérienne ont étéensemencés sur des géloses contenant de l'ampicilline (100µg/ml). Les géloses ont été incubées une nuit à 37°C.

Le lendemain, plusieurs clones bactériens obtenus sont analysés par PCR afin de sélectionner ceux ayant incorporé le plasmide contenant l'insert. Les amorces utilisées à cette étape (V1F/V1R) s'hybrident dans l'intron 3 du vecteur pTB2 de part et d'autre du site de restriction *NdeI*. La

composition du mélange réactionnel et les conditions de PCR utilisées sont présentées dans les Tableau 23 et Tableau 24 :

Réactifs	Volume (µl)
Tampon Qbiogene with MgCl <sub>2</sub> 10X	2,5
Amorces V1F [10pM]	0,625
Amorces V1R [10pM]	0,625
dNTP [2,5 mM de chaque dNTP]	1
Taq DNA polymerase MPbiomedicals [5U/µl]	0,25
Eau	Qsp 25
Clone bactérien	1 clone

**Tableau 23 : Composition du mélange réactionnel pour la vérification des clones bactériens.**

	Température (°C)	Temps	Nombre de cycles
Activation	95	5 min	1
Dénaturation	95	20 s	
Hybridation	55	20 s	35
Elongation	72	40 s	
Elongation finale	72	4 min	1

**Tableau 24 : Conditions de PCR pour la vérification des clones bactériens.**

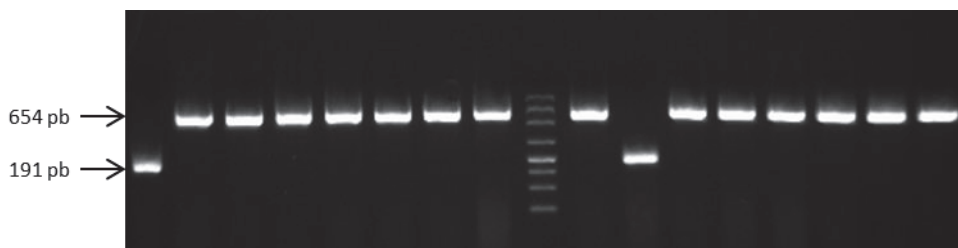
Dans le cas du minigène contenant le premier exon du gène *F8*, les amorces utilisées s'hybrident au niveau du promoteur (amorce sens pTBprom\_F) et dans l'exon 2 du minigène (amorce antisens αglobuline Ex2 rev).

Séquences des amorces utilisées pour la sélection des clones bactériens:

V1F : AGCATTGCTGATACTTACTTCAGA  
V1R : CACAGCAAAGAGAAACATAGCAG  
pTBprom\_F : AAAAAGCTCCAAGCTCTCG  
αglobuline Ex2 rev : GTGGCCCTTAACCTGGGCAGAG

Les clones bactériens sélectionnés ont été ajoutés au mélange réactionnel de PCR puis repiqués dans 1 ml de milieu LB liquide pendant 8h à 37°C sous agitation. Les produits de PCR ont été analysés par migration sur gel d'agarose. Une bande de 191 pb était visible pour les clones bactériens ayant assimilé le vecteur vide et une bande de plus grande taille (191 pb + taille de l'insert) était visible pour ceux ayant assimilé le vecteur contenant l'insert (dans le cas de l'exon 1, la bande correspondant au vecteur vide est de 204 pb) (Figure 30).





**Figure 30 : Exemple de migration des produits de PCR issus de la vérification des clones bactériens transformés avec le minigène contenant l'exon 5 du gène *F8*.** Les bandes de 191 pb correspondent à des vecteurs n'ayant pas assimilé l'insert lors de la réaction In-fusion. Les bandes de 654 pb (191 pb du vecteur + 220 pb IVS4 + 69 pb exon 5 + 174 pb IVS5) correspondent à des vecteurs ayant assimilé l'insert lors de la réaction In-fusion. Les clones bactériens correspondant aux bandes de 654 pb ont été sélectionnés pour la production du minigène contenant l'exon 5 du gène *F8*.

Ensuite, 150  $\mu$ l de suspension bactérienne correspondant aux clones sélectionnés ont été dilués au 1/1000 dans 150 ml de milieu LB et incubés une nuit à 37°C sous agitation. Le lendemain, les bactéries ont été culotées (centrifugation 3000g pendant 20 minutes) puis les plasmides purifiés par chromatographie échangeuse d'anions (Qiagen Plasmid Maxi Kit, Qiagen, Venlo, pays-bas) en suivant les indications du fournisseur. Les éluats finaux ont été dosés par spectrophotométrie d'absorption à 260 nm.

La région contenant l'insert a été séquencée pour tous les plasmides afin de vérifier la présence des mutations d'intérêt ainsi que l'absence d'autres mutations. Dans une première étape, la région d'intérêt a été amplifiée à l'aide du couple d'amorce V1R/V1F comme précédemment. Pour la réaction de séquence, le kit BigDye Terminator Cycle sequencing (Life technologies) a été utilisé en suivant les recommandations du fournisseur (mélange réactionnel : BigDye 0,5 $\mu$ l, tampon 5X 1,75 $\mu$ l, amorces V1R ou V1F 5pmol, produit de PCR 1 $\mu$ l, eau qsp 8 $\mu$ l). Les fragments obtenus ont été purifiés (BigDye Xterminator purification Kit, Life technologies) puis injectés dans un séquenceur capillaire ABI3130 (Applied Biosystem). Les séquences ont été lues à l'aide du logiciel Seqscape v3.0 (Life technologies). Le détail des différents plasmides produits est présenté dans le Tableau 25.

Exon du gène <i>F8</i> étudié	Détails de la construction du minigène
1	302 pb exon 1 + 102 pb IVS1
2	228 pb IVS1 + 122 pb exon 2 + 219 pb IVS2
3	178 pb IVS2 + 123 pb exon 3 + 256 pb IVS3
4	202 pb IVS3 + 213 pb exon 4 + 220 pb IVS4
5	220 pb IVS4 + 69 pb exon 5 + 174 pb IVS5
6	234 pb IVS5 + 117 pb exon 6 + 141 pb IVS6
7	243 pb IVS6 + 222 pb exon 7 + 213 pb IVS7
8	430 pb IVS7 + 262 pb exon 8 + 184 pb IVS8
11	140 pb IVS10 + 215 pb exon 11 + 156 pb IVS 11
13	202 pb IVS12 + 210 pb exon 13 + 122 pb IVS13
15	100 pb IVS14 + 154 pb exon 15 + 324 pb IVS15
16	227 pb IVS15 + 213 pb exon 11 + 207 pb IVS16
18	108 pb IVS17 + 183 pb exon 18 + 185 pb IVS18
19	222 pb IVS18 + 117 pb exon 19 + 275 pb IVS19

**Tableau 25 : description des différents minigènes produits dans cette étude.**

#### d) Transfection cellulaire

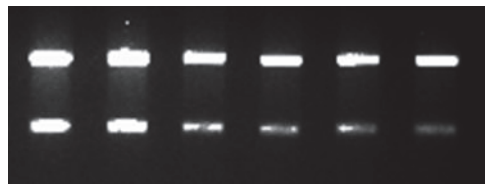
Deux lignées cellulaires ont été utilisées dans cette étude. Les cellules HeLa et Huh7 ont été cultivées dans des flasques de 25 cm<sup>2</sup> dans du milieu DMEM-F12 (Life technologies) en présence de 10 % de sérum de veau fœtal (SVF, Thermofischer, Waltham, MA, États-Unis) et de 1% de pénicilline-streptomycine (Life technologies) à 37°C sous 5% de CO<sub>2</sub>.

Les cellules ont été traitées par la trypsine 1X (Life technologies) lorsqu'elles étaient à 70% de confluence environ. Après comptage, 100 000 cellules ont été réparties dans chaque puits d'une plaque six puits et le volume a été complété à 2ml avec du milieu DMEM-F12/10%SVF. Les cellules ont été ensuite incubées une nuit à 37°C.

Le lendemain, le milieu de culture a été éliminé puis les cellules ont été rincées au DPBS 1X. Les cellules ont été transfectées transitoirement à l'aide du réactif X-tremeGENE9 DNA (Roche molecular diagnostics, Meylan, France). Dans un tube eppendorf ont été mélangés 100 µl de milieu DMEM-F12 sans SVF, 1µg de plasmide et 3 µl de transfectant. Après 15 minutes d'incubation à température ambiante, le mélange réactionnel a été déposé goutte à goutte dans un puits. Le volume a ensuite été complété à 2 ml avec du milieu DMEM-F12/SVF10%. Pour chaque plasmide (muté ou témoin), les transfection ont été réalisées en triplicat.

e) Extraction des ARNm totaux

Après 48h d'incubation à 37°C sous 5% de CO<sub>2</sub>, les surnageants cellulaires ont été éliminés, les tapis cellulaires ont été lavés deux fois avec du DPBS 1X puis lysés. Les ARNm totaux ont été extraits des lysats cellulaires à l'aide du kit RNAqueous-4PCR (Life technologies) en suivant les recommandations du fournisseur. Les ARNm purifiés ont ensuite été élués dans 40µl de tampon TE préchauffé à 70°C, traités à la DNase et dosés. La qualité des ARNm totaux extraits a été vérifiée par migration sur gel d'agarose 1% (Figure 31) et par calcul du ratio DO<sub>260nm</sub>/DO<sub>280nm</sub> (un ratio > 2 est un marqueur de bonne qualité des ARNm).



**Figure 31 : Exemple de migration sur gel d'agarose 1% des ARNm totaux purifiés extraits à partir de lysats de cellules HeLa.** La présence de deux bandes bien distinctes correspondant aux ARN ribosomique 28S (5kb) et 18S (1,9kb) est un marqueur de bonne qualité des ARNm extraits. La présence d'une trainée (smear) serait un signe de dégradation des ARNm.

f) Etude de l'épissage de l'exon d'intérêt

La réaction de RT-PCR a été réalisée en utilisant le kit transcriptor High Fidelity cDNA Synthesis Sample (Roche Molecular Diagnostics). Dans une première étape, 250 ng d'ARNm totaux ont été dénaturés à 65°C pendant 10 minutes en présence d'héxanucléotides randomisés (1µl à 600µM). La réaction de RT-PCR a été réalisée dans une seconde étape. La composition du milieu réactionnel est présentée dans le Tableau 26 :

Réactifs	Volume (µl)
Tampon RT transcriptor Reaction 5X	4
RNase inhibitor (40U/µl)	0,5
DTT 0,1M	1
Mélange ARNt + hexanucléotide	10,4µl
dNTP (10mM chacun)	2
Reverse transcriptase (20U/µl)	1,1

**Tableau 26 : Composition du mélange réactionnel pour la réaction de RT-PCR.** DTT : dithiothreitol.

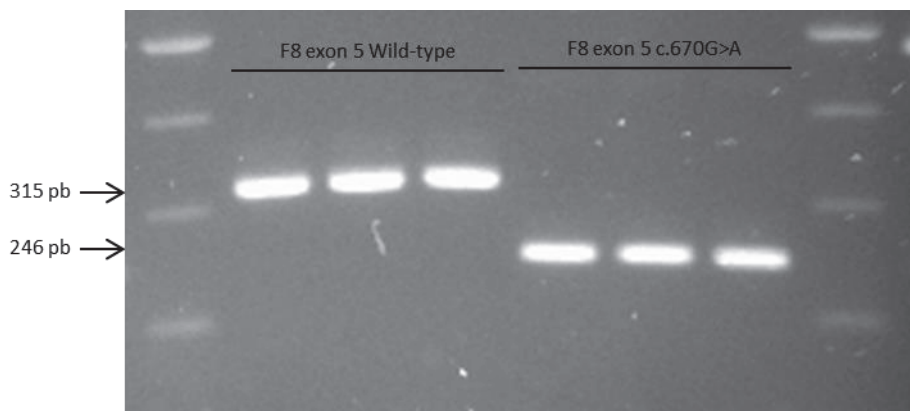
La réaction de RT-PCR est décomposée en deux étapes, une phase d'élongation à 45°C pendant 30 minutes suivie d'une phase d'inactivation de la transcriptase inverse à 85°C pendant 5 minutes.

Dans une dernière étape, les ADNc obtenus ont été amplifiés par PCR en utilisant des amorces spécifiques s'hybridant au niveau de la jonction des exons 2 et 3 (amorce sens  $\alpha$ 23G) et au niveau de l'exon 4 du minigène (amorce antisens  $\beta$ 2F). La position des amorces est indiquée au niveau de la Figure 29. Les amorces utilisées sont spécifiques des transcrits du minigène car elles s'hybrident sur des exons chimériques  $\alpha$ globine/fibronectine. Dans le cas du minigène contenant le premier exon du gène *F8*, les amorces utilisées sont les mêmes que celles utilisées pour la sélection des clones bactériens (pTBprom\_F/ $\alpha$ globuline Ex2 rev). La composition du mélange réactionnel et les conditions de PCR sont identiques à celle utilisées pour la sélection des clones bactériens (Tableau 23 et Tableau 24).

Les séquences des amorces utilisées l'amplification des ADNc sont:

$\alpha$ 23G : CAACTTCAAGCTCCTAAGCCACTG  
 $\beta$ 2F : TGATTTAACCAACTTCCTGGTGACC

Les produits de PCR ont été étudiés par migration sur gel d'agarose à 2% à 70V pendant 90 minutes et par séquençage (amorces de séquençage :  $\alpha$ 23G et  $\beta$ 2F). Les bandes attendues sont de 246 pb dans le cas d'un saut d'exon et de 246+taille de l'exon dans le cas d'un épissage correct de l'exon étudié (Figure 32).



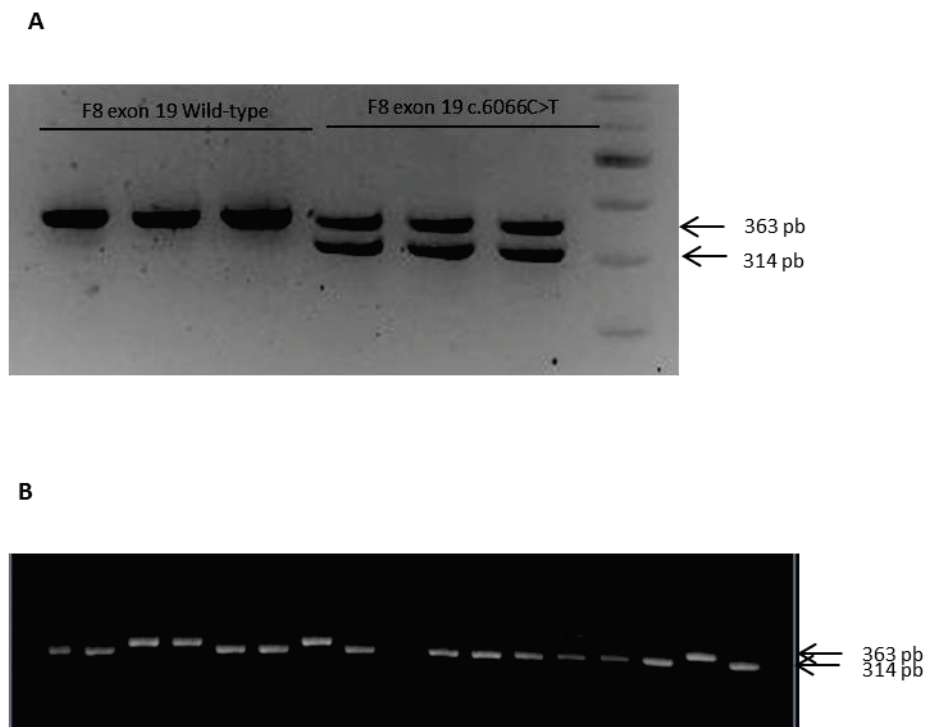
**Figure 32 : Exemple d'analyse des transcrits produits par des cellules HeLa après transfection par les minigènes *F8* exon 5 Wild-type (exon normal) et *F8* exon 5 c.670G>A.** La bande de 315 pb correspond à l'épissage correct de l'exon 5 du gène *F8*, la bande de 246 pb correspond au saut complet de cet exon. Dans cet exemple, la mutation c.670G>A est responsable du saut de l'exon 5.

Dans le cas de la présence de plusieurs produits de PCR, une étape de sous-clonage dans un plasmide TOPO a été réalisée pour isoler chacune des bandes avant de les séquencer (TOPO TA cloning kit, Life technologies). Le mélange réactionnel est présenté dans le Tableau 27.

Réactifs	Volume ( $\mu$ l)
PCR récente	1
Salt solution	1
Eau	2
TOPO Vector	1

**Tableau 27 : Composition du mélange réactionnel pour la réaction TOPO.**

Le mélange réactionnel a été incubé 20 minutes à température ambiante puis des bactéries compétentes DH5 $\alpha$  ont été transformées par 2 $\mu$ l du mélange réactionnel précédent. Après 1 heure d'incubation à 37°C dans du milieu S.O.C, 100  $\mu$ l de la suspension bactérienne ont été ensemencés sur des géloses à l'ampicilline. Après une nuit à 37°C, 15 à 30 clones bactériens ont été analysés par PCR (amorces utilisées :  $\alpha$ 23G et  $\beta$ 2F). Après migration en gel d'agarose, des fragments de différentes tailles ont été séquencés (Figure 33).



**Figure 33 : Exemple de clonage TOPO réalisé pour l'étude de la mutation c.6066C>T.** (A) Analyse des transcrits produits par des cellules HeLa après transfection par les minigènes *F8* exon 19 Wild-type (exon normal) et *F8* exon 19 c.6066C>T. La bande de 363 pb correspond à l'épissage correct de l'exon 19 du gène *F8*. Dans le cas du minigène porteur de la mutation, deux bandes sont observées, une bande de 363 pb correspondant à l'épissage normal de l'exon 19 et une autre bande de 314 pb. (B) un clonage TOPO est réalisé pour isoler les produits de PCR issus du minigène c.6066C>T. 16 clones bactériens sont analysés. Les deux bandes de 363 et 314 pb sont correctement isolées permettant leur séquençage. Le

séquençage révélera l'utilisation d'un site cryptique donneur d'épissage c.6064 entraînant la perte des 51 derniers nucléotides de l'exon 19.

#### 4. Expression cellulaire de variants du FVIII dans le cas des mutations faux-sens

Certaines mutations exoniques incluses dans cette étude sont des mutations faux-sens. Afin d'évaluer l'impact du changement d'acide aminé sur la synthèse et l'activité du FVIII, des expériences d'expression cellulaire de variant du FVIII ont été réalisées.

##### a) Description du plasmide utilisé

Le plasmide pcDNA3-F8<sub>1+13</sub> contenant l'ADNc du gène *F8* a été utilisé (225). Les caractéristiques de cette construction sont les suivantes :

- ADNc du gène *F8* dépourvu de la région de l'exon 14 codant le domaine B (l'exon 14 est réduit aux parties codant les régions acides a2 et a3 séparées par un linker de 12 nucléotides) est sous le contrôle du promoteur du Cytomégalo virus.
- La présence d'une séquence nucléotidique appartenant à l'intron 1 du gène *F9* insérée entre les exons 1 et 2 et les exons 13 et 14 du gène *F8*, permet d'augmenter significativement la production de FVIII *in vitro* (225).
- Présence du gène de résistance à l'ampicilline (*ampR*) qui permet la sélection des clones bactériens ayant assimilés le plasmide.

#### **Production des plasmides mutants**

Les substitutions nucléotidiques étudiées ont été insérées dans le vecteur pcDNA3-F8<sub>1+13</sub> par mutagenèse dirigée en utilisant le kit QuickChange II site-Directed Mutagenesis (Stratagene, La Jolla, Ca Etats-Unis) en suivant les recommandations du fournisseur. La liste des amorces de mutagenèse utilisées est présentée en annexe de cette section. La composition du mélange réactionnel ainsi que les conditions de réaction sont indiquées dans les Tableau 28 et Tableau 29:

Réactifs	Volume (μl)
QuickChange Multi reaction buffer 10 X	5
pcDNA3-F8 <sub>1+13</sub>	80 ng
Amorce sens	10 pmol
Amorce antisens	10 pmol
dNTP Mix*	1
Pfu Ultra (2,5U/μl)	1
Eau	Qsp 50

**Tableau 28 : Composition du mélange réactionnel pour la réaction de mutagenèse dirigée.** \* Composition non indiquée par le fournisseur.

	Température (°C)	Temps	Nombre de cycles
Dénaturation	98	20 s	
Hybridation	55	10 s	20
Elongation	68	20 min	

**Tableau 29 : Condition de PCR pour la réaction de mutagenèse dirigée.**

Afin de dégrader le plasmide initial non muté en fin de mutagenèse dirigée, 1μl d'enzyme *DpnI* à 10U/μl (enzyme de restriction digérant spécifiquement l'ADN méthylé, séquence cible : 5'-Gm<sup>6</sup>ATC-3') a été ajouté au mélange réactionnel et incubé 1h30 à 37°C.

Les plasmides mutés ont ensuite été produits de la même manière que les vecteurs minigènes. Des bactéries DH5α ont été transformées par choc thermique avec 5 μl du mélange réactionnel de mutagenèse et cultivé en présence d'ampicilline. La purification des plasmides a été réalisée par chromatographie d'échange d'anions à l'aide du kit Qiagen Plasmid MaxiKit. Chaque plasmide produit a été vérifié par séquençage (liste des amorces en annexe de cette section).

### Transfection cellulaire

Les cellules COS-1 ont été utilisées pour l'expression de variant du FVIII. Les cellules ont été cultivées dans des flasques de 75 cm<sup>2</sup> dans du milieu IMDM (Life technologies) en présence de 10% de SVF et 1% de pénicilline-streptomycine. Les cellules ont été traitées par la trypsine 1X lorsqu'elles étaient à 70% de confluence environ. Après comptage, 500 000 cellules ont été réparties dans chaque puits d'une plaque six puits et le volume a été complété à 2ml avec du milieu IMDM/10%SVF. Les cellules ont été ensuite incubées une nuit à 37°C.

Le lendemain, le milieu de culture a été éliminé puis les cellules ont été rincées au DPBS 1X. Les cellules ont été transfectées transitoirement à l'aide du réactif X-tremeGENE9. Dans un tube eppendorf ont été mélangés 100 μl de milieu IMDM sans SVF, 2μg de plasmide et 5 μl de

transfectant. Après 15 minutes d'incubation à température ambiante, le mélange réactionnel a été déposé goutte à goutte dans un puits. Le volume a ensuite été complété à 2 ml avec du milieu IMDM/SVF10%.

Après 36h d'incubation à 37°C sous 5% de CO<sub>2</sub>, le surnageant cellulaire a été éliminé, les tapis cellulaires lavés deux fois avec du DPBS1X. Cinq cents microlitres de milieu IMDM enrichi par 1% d'albumine bovine (BSA) ont été ajoutés dans chaque puits. Les cellules ont été incubées une nuit à 37°C sous 5% de CO<sub>2</sub>. Le lendemain les surnageants de culture cellulaire ont été recueillis, centrifugés (3000g, 15 min), aliquotés et congelés à -80°C. Après deux lavages DPBS1X, les tapis cellulaires ont été lysés par ajout de 144µl de tampon de lyse (HEPES pH7,5 50mM, KCl 0,1M, MgCl<sub>2</sub> 2mM, Triton X-100 0,5%) et de 6 µl d'une solution de d'inhibiteur de protéases (Complete Mini, Roche, Bâle, Suisse). Après une incubation de 20 minutes dans la glace, les puits ont été grattés puis 150 µl de DPBS 1X ont été ajoutés. Les lysats cellulaires ont été centrifugés à 14000g pendant 5 minutes puis aliquotés et congelés à -80°C.

Chaque transfection a été réalisée en duplicat. Un contrôle négatif (cellules non transfectées) et un contrôle normal (cellules transfectées avec le plasmide wild-type) ont été insérés dans chaque série de transfection. Trois séries de transfusions indépendantes ont été réalisées pour chaque variant étudié.

#### b) Dosage du FVIII produit

Les dosages d'activité du FVIII dans les surnageants cellulaires ont été réalisés par technique coagulométrique (FVIII:C) et chromogénique (FVIII:CR). Les mesures coagulométriques ont été réalisées sur un ACL top 750 (Instrument Laboratory, Bedford, Ma, Etat-Unis) avec la céphaline Synthasil (Instrument Laboratory). Les dosages chromogéniques ont été réalisés à l'aide du kit Biophen FVIII:C kit (Hyphen Biomed TM, Neuville-sur Oise, France). Les dosages antigéniques (FVIII:Ag) ont été réalisés par méthode ELISA (Assera-chrom FVIII kit; Stago, Asnières sur seine, France). Tous les résultats ont été exprimés en pourcentage du FVIII-WT. Les résultats sont présentés sous la forme moyenne +/- écart type.



## 4. Résultats

### 1. Résultats globaux

Des tests fonctionnels *in vitro* ont été réalisés afin de déterminer l'impact de 26 variations nucléotidiques sur l'épissage du gène *F8*. Ces variations sont localisées soit au niveau d'un site donneur (n = 18) soit au niveau d'un site accepteur (n=8) d'épissage (Tableau 20). Pour chacune de ces variations, une combinaison de tests *in silico* et *in vitro* a été utilisée afin de déterminer leur impact sur l'épissage du gène *F8*. Les résultats détaillés de ces investigations sont présentés dans les Tableau 30 et Tableau 31. Pour 24/26 variants étudiés, aucune différence n'a été observée lors de l'analyse *in vitro* entre les deux lignées cellulaires utilisées (HeLa et Huh7). Des anomalies d'épissage partielles ou totales ont été démontrées pour 21/26 des variations nucléotidiques étudiées. La pathogénicité des 5 variants restants est incertaine. Des résultats concordants entre l'analyse *in silico* et *in vitro* ont été obtenus pour 16/18 (89%) variants des sites 5'ss et pour 2/8 (25%) des variants des sites 3'ss.

### 2. Impact sur l'épissage du gène *F8* des variations nucléotidiques situées au niveau d'un site donneur d'épissage.

Huit des 18 variants situés au niveau d'un site donneur d'épissage sont exoniques. Parmi ces 8 variants, 4 sont localisés au niveau du dernier nucléotide d'un exon, entraînant le remplacement du nucléotide consensus G à la position -1 du site donneur par un autre nucléotide (c.670G>A, c.1009G>A, c.1752G>A et c.5586G>A). L'analyse *in silico* est en faveur d'un effet délétère sur l'épissage de ces substitutions. En effet une diminution significative de la force des sites 5'ss naturel a été observée dans chaque cas (Tableau 30). L'étude par minigène de ces variations a montré que les mutations c.670G>A, c.1752G>A et c.5586G>A entraînaient :

- un saut d'exon partiel (coexistence des transcrits normaux et aberrants) : c.1752G>A, le saut de l'exon 11 est responsable d'un décalage du cadre de lecture, p.Gly513Aspfs\*15
- un saut d'exon total (absence de transcrit normal) : les mutations c.670G>A et c.5586G>A entraînent respectivement le saut de l'exon 5 (p.Gly224\_Gly263delinsArg) et de l'exon 15

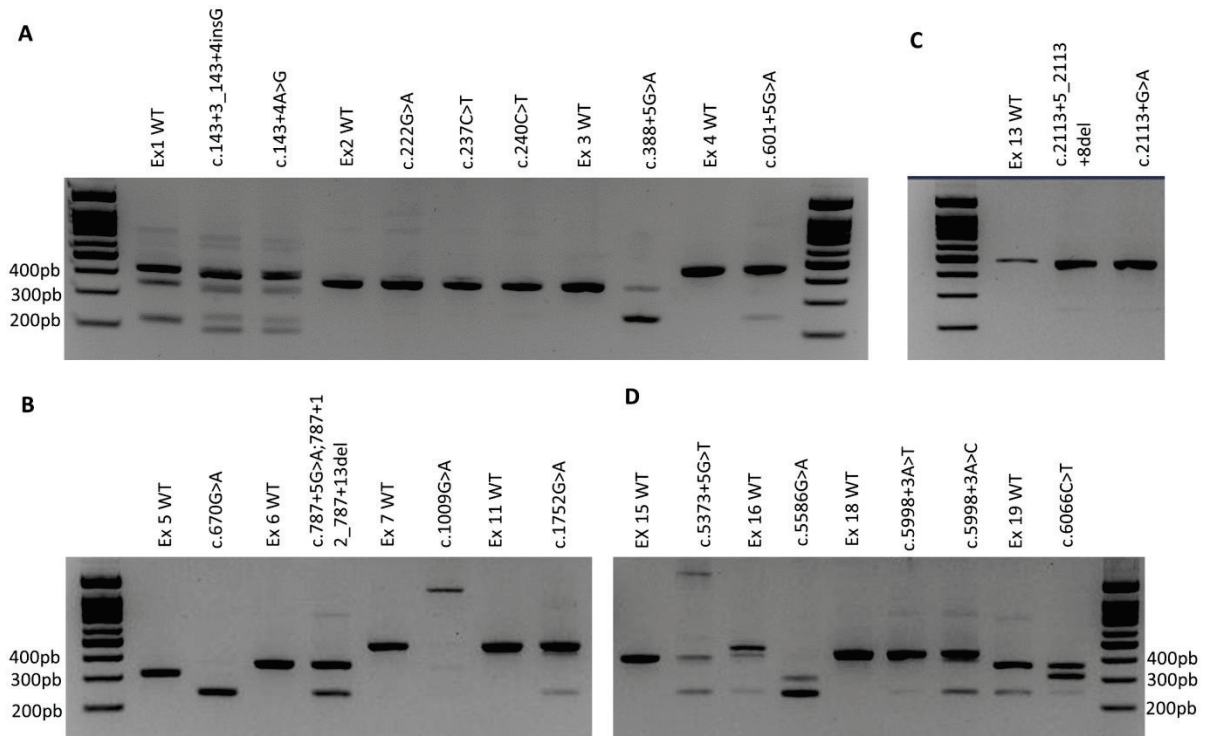
(p.Val1792\_Leu1862del), le cadre de lecture est conservé dans les deux cas (Figure 34 Figure 35 B et D).

Dans le cas de la variation c.1009G>A, un fragment majoritaire de grande taille (1123 bp) composé de l'exon 7 et d'une partie de l'intron 7 non épissé a été observé lors des analyses minigènes (Figure 34 Figure 35 B).

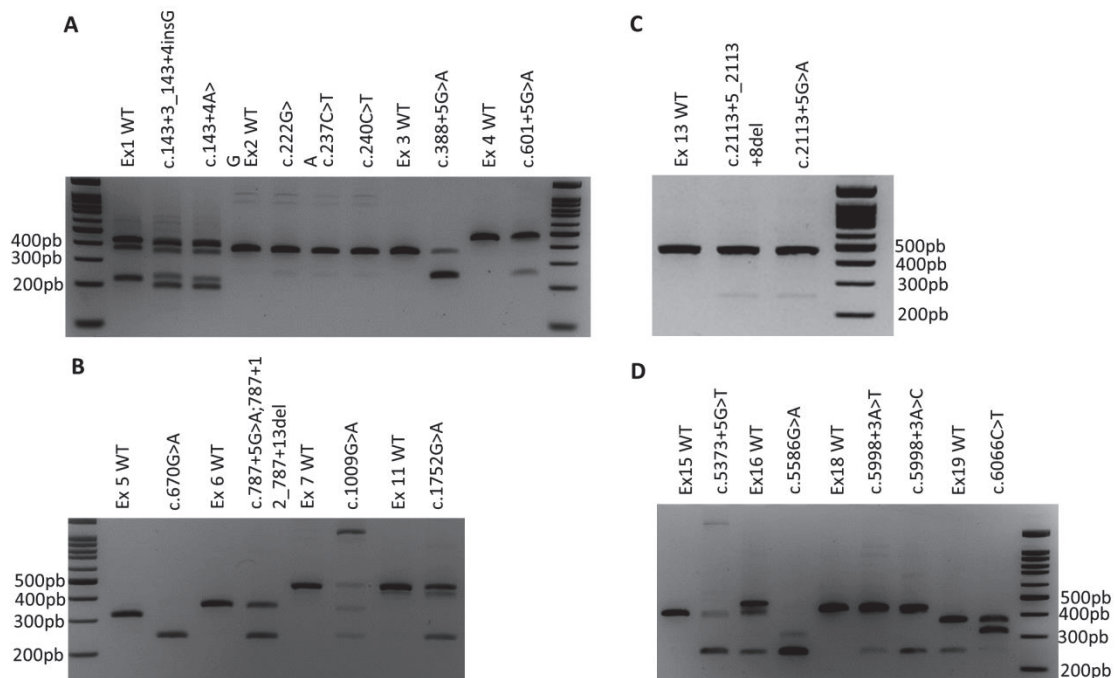
Comme la présence de traces de transcrits correctement épissés ne pouvait être exclue, l'impact des mutations faux-sens c.670G>A (p.Gly224Arg) et c.1009G>A (p.Asp337Asn) sur la biosynthèse et l'activité du FVIII a été évalué par expression cellulaire de variant du FVIII. Aucun impact délétère n'a été mis en évidence pour la mutation p.Gly224Arg alors qu'une diminution significative à la fois des activités FVIII:C et FVIII:CR et du FVIII:Ag dans les surnageants cellulaires a été observée pour la mutation p.Asp337Asn (FVIII:C = 29%, FVIII:CR = 27% et FVIII:Ag = 33%) (Figure 36).

Variations	Analyse <i>In silico</i>		Analyse <i>in vitro</i> par minigène
	Force du site physiologique > Force du site muté MaxEntScan	Human splicing finder	
Variations au niveau des sites donneurs d'épissage			
c.143+3_143+4insG	9.06 > 6.33 (-30.2%) Cryptique c.121: 9.10	84.95 > 76 (-10.3%) Cryptique c.121: 89.93	Utilisation du site cryptique c.121
c.143+4A>G	9.06 > 6.38 (-29.7%) Cryptique c.121 : 9.10	84.95 > 76.61 (-9.8%) Cryptique c.121 : 89.93	Utilisation du site cryptique c.121
c.222G>A	5.68 > 5.68 (0%)	78.91 > 78.91 (0%)	Epissage normal de l'exon 2
c.237C>T	5.68 > 5.68 (0%)	78.91 > 78.91 (0%)	Epissage normal de l'exon 2
c.240C>T	5.68 > 5.68 (0%)	78.91 > 78.91 (0%)	Epissage normal de l'exon 2
c.388+5G>A	9.25 > 2.65 (-71.4%)	93.03 > 80.06 (-13.1%)	Epissage normal et saut de l'exon 3
c.601+5G>A	11 > 8.88 (-19.2%)	99.04 > 86.88 (-12.3%)	Epissage normal et saut de l'exon 4
c.670G>A	8.54 > 4.24 (-50.3%)	94.02 > 83.44 (-11.3%)	Saut de l'exon 3
c.787+5G>A and c.787+12_787+13del	9.8 > 7.88 5 (-19.6%)	91.19 > 79.03 (-13.3%)	Epissage normal et saut de l'exon 6
c.1009G>A	5.05 > non détecté; Cryptique c.888: 3.38	79.64 > 69.06 (-13.3%) Cryptique c.888: 81.29	Epissage normal et saut de l'exon 7
c.1752G>A (p.Gln584Gln)	10.67 > 9.10 (-14.8%)	96.84 > 88.26 (-10.7%)	Epissage normal et saut de l'exon 11
c.2113+5_2113+8del	8.02 > 3.49 (-56.5%)	94.97 > 82.66 (-13%)	Epissage normal et saut de l'exon 13 (uniquement pour le Huh7)
c.2113+5G>A	8.02 > 4.70 (-41.4%)	94.97 > 82.81 (-12.8%)	Epissage normal et saut de l'exon 13 (uniquement pour le Huh7)
c.5373+5G>T	10.13 > 3.18 (-68.6%)	93.14 > 80.82 (-13.2%)	Epissage normal et saut de l'exon 15
c.5586G>A	8.69 > 5.20 (-40.2%)	93.07 > 82.50 (-11.4%)	Epissage normal et saut de l'exon 16
c.5998+3A>T	9.46 > 5.11 (-46%)	94.16 > 82.37 (-12.5%)	Epissage normal et saut de l'exon 18
c.5998+3A>C	9.46 > 5.29 (-46.5%)	94.16 > 83.11 (-13.4%)	Epissage normal et saut de l'exon 18
c.6066C>T (p.Gly2003Gly)	6.97 Cryptique c.6064 : non détecté > 7.43	82.56 Cryptique c.6064 : non détecté > 89.93	Epissage normal et utilisation du site cryptique c.6064

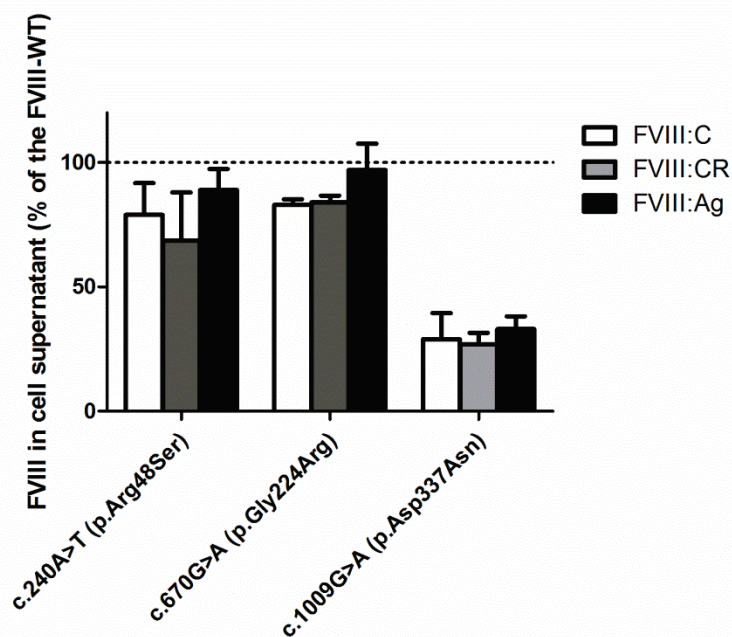
**Tableau 30 : Résultats des analyses *in silico* et *in vitro* obtenus pour les variants situés au niveau d'un site donneur d'épissage.** Pour l'analyse *in silico*, un variant a été considéré comme affectant l'épissage si les prédictions indiquaient une différence d'au moins 15% pour MaxEntScan et d'au moins 5% pour HSF entre la force du site physiologique et celle du site muté. Dans le cas des sites *de novo* d'épissage, le variant a été considéré comme délétère si la force du site *de novo* était égale ou supérieure à celle du site naturel d'épissage. Pour l'analyse *in vitro*, un variant a été considéré comme délétère si un épissage aberrant était observé dans les deux modèles cellulaires. Les variants grisés correspondent à des discordances entre l'analyse *in silico* et *in vitro*.



**Figure 34 : Résultats de l'analyse *in vitro* effectuée après transfection de cellules HeLa pour les 18 variations situées au niveau d'un site donneur d'épissage.** Migration sur gel d'agarose des produits de RT-PCR obtenus pour les variations affectant les exons 1 à 4 (A), 5 à 11 (B), 13 (C) et 15 à 19 (D). Un saut d'exon correspond à une bande de 246 pb dans notre modèle de minigène. WT : Wild Type.



**Figure 35 : Résultats de l'analyse *in vitro* effectuée après transfection de cellules Huh7 pour les 18 variations situées au niveau d'un site donneur d'épissage.** Migration sur gel d'agarose des produits de RT-PCR obtenus pour les variations affectant les exons 1 à 4 (A), 5 à 11 (B), 13 (C) et 15 à 19 (D). Un saut d'exon correspond à une bande de 246 pb dans notre modèle de minigène. WT : Wild Type.



**Figure 36 : Activité du FVIII mesurée par méthode chromométrique (FVIII:C) et chromogénique (FVIII:CR) et FVIII antigène (FVIII:Ag) mesuré dans les surnageants de culture cellulaire 48h après la transfection des cellules COS-1. Les données sont exprimées en pourcentage par rapport à un FVIII Wild-type exprimé dans les mêmes conditions. Chaque expression a été réalisée en doublon et répétée trois fois de manière indépendante (n=3).**

Parmi les 4 substitutions exoniques situées près d'un site donneur d'épissage, 3 étaient des variations synonymes localisées dans l'exon 2 du gène *F8* (c.222G>A, c.237C>T, c.240C>T). L'analyse *in silico* n'a prédit aucun impact sur la force du 5'ss de l'exon 2 (Tableau 30) et seul le transcrit correspondant à l'épissage correct de l'exon 2 (bande de 368 pb) a été observé lors de l'étude par minigène de ces trois variants (Figure 34 Figure 35 A).

Pour la dernière substitution exonique étudiée dans ce travail, c.6066C>T, l'étude *in silico* était en faveur de la création d'un nouveau site donneur d'épissage à la position c.6064 de force supérieure à la force du site naturel (résultat MaxEntScan : score du site *de novo* 7,43 vs 6,43 pour le site naturel). L'étude par minigène de cette mutation confirme la reconnaissance par le splicéosome du site *de novo* en c.6064 entraînant la co-existence de deux transcrits, l'un de 363 pb correspondant à l'épissage correct de l'exon 19 et l'autre de 312 pb correspondant à l'utilisation du site *de novo* en c.6064 (perte de 51 dernières paires de bases de l'exon 19, conséquence protéique : p.Gly2022\_Lys2079delinsGlu) (Figure 34 Figure 35 D).

Les dix autres variants affectant un site donneur d'épissage sont introniques : 8 sont des substitutions situées en positions +3 (c.5998+3A>T et c.5998+3A>C), +4 (c.143+4A>G), et +5 (c.388+5G>A, c.601+5G>A, c.2113+5G>A, c.5373+5G>t et c.[787+5G>A ;787+12\_787+13del) et les deux derniers sont une petite insertion (c.143+3\_143+4insG) et une petite délétion (c.2133+5\_2113+8del). L'analyse *in silico* a montré que toutes ces variations nucléotidiques étaient responsables d'une diminution significative de la force des sites 5'ss naturels (Tableau 30). L'impact sur l'épissage de ces variants a été étudié *in vitro* dans deux modèles cellulaires (HeLa et Huh7).

Les variants c.143+4A>G et c.143+3\_+4insG sont localisés au niveau du site donneur de l'intron 1 du gène *F8*. Trois transcrits différents ont été observés dans les cellules transfectées (cellules HeLa et Huh7) par le minigène correspondant à l'exon 1 WT. Après sous-clonage de ces produits de PCR et séquençage, il a été déterminé que le transcrit majoritaire de 418 pb correspond à l'épissage correct de l'exon 1, les deux autres transcrits sont des artéfacts de la technique. Les variations nucléotidiques c.143+4A>G et c.143+3\_+4insG sont responsables de la production d'un transcrit majoritaire plus court (396 pb) (Figure 34 Figure 35 D). Le séquençage de ce fragment de 396 pb a montré que ces deux mutations étaient responsables de l'activation d'un site donneur cryptique d'épissage en c.121 entraînant la perte des 22 derniers nucléotides de l'exon 1 (l'utilisation de ce site cryptique d'épissage entraîne un décalage du cadre de lecture au niveau protéique, p.Gly41Aspfs24\*). Le transcrit normal de 418 pb n'a pas été trouvé pour ces deux mutations.

Les études *in vitro* des variants c.388+5G>A, 601+5G>A, c.[787+5G>A;787+12\_13del], c.5998+3A>T and c.5998+3A>C dans les deux lignées cellulaires ont montré dans chaque cas la coexistence de transcrits normaux et de transcrits aberrants correspondant à un saut d'exon (bande de 246 pb) (Figure 34 Figure 35 A, B, D). Ces six mutations sont donc responsables d'un saut partiel d'exon.

Enfin, les analyses *in vitro* des deux variations affectant le site donneur de l'exon 13 du gène *F8* (c.2113+5G>A et c.2133+5\_2113+8del) sont difficiles d'interprétation du fait de résultats légèrement différents entre les deux lignées cellulaires testées. Seule la bande correspondant à l'épissage normal

de l'exon 13 a été observée dans les cellules HeLa pour ces deux variants (bande de 456 pb) alors qu'en plus de l'épissage normal, des traces de saut de l'exon 13 (bande de 246 pb) ont été observées dans les cellules Huh7 (Figure 34 Figure 35 C).

### 3. Impact sur l'épissage du gène *F8* des variations nucléotidiques situées au niveau d'un site accepteur d'épissage.

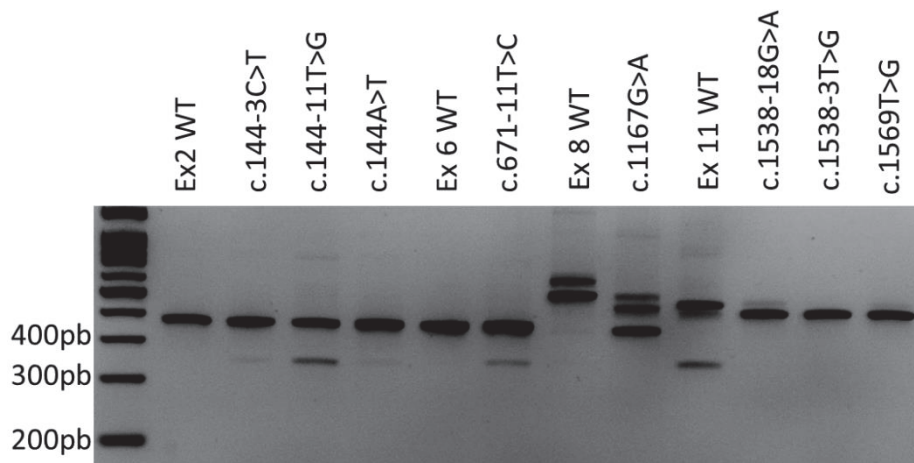
Trois des huit substitutions localisées au niveau d'un site accepteur du gène *F8* sont exoniques (c.144A>T, c.1167A>G et c.1569G>T). Le variant c.144A>T affecte le premier nucléotide de l'exon 2. Seul MaxEntScan était en faveur d'un effet délétère sur l'épissage de cette substitution (Tableau 31). L'analyse fonctionnelle de cette anomalie a montré l'existence de deux transcrits, l'un (très minoritaire) correspondant à un saut de l'exon 2 (246 pb) et entraînant l'apparition d'un codon stop (p.Phe49serfs\*9) et l'autre (majoritaire) correspondant au transcrit normal (Figure 37 et Figure 38). Comme la substitution c.144A>T entraîne un changement d'acide aminé (p.Arg48Ser), l'impact de ce variant sur la production et l'activité du FVIII a été étudié par expression cellulaire de variant dans des cellules COS-1. Des résultats d'activités et de FVIII:Ag dans les surnageants cellulaires comparables à ceux obtenus pour le FVIII-WT ont été observés (Figure 36).

Variations	Analyse <i>In silico</i> Force du site physiologique > Force du site muté		Analyse <i>in vitro</i>
	MaxEntScan	Human splicing finder	
Variations au niveau des sites accepteurs d'épissage			
c.144-3C>T	11.8 > 10.21 (-8.7%)	90.62 > 82.93 (-8.5%)	Epissage normal et saut de l'exon 2
c.144-11T>G	11.8 > 8.97 (-19.8%)	90.62 > 88.96 (-4%)	Epissage normal et saut de l'exon 2
c.144>A>T	11.8 > 8.41 (-15.9%)	90.62 > 89.45 (-1.3%)	Epissage normal et saut de l'exon 2
c.671-11T>C	11.76 > 10.74 (-8.7%)	92.8 > 90.95 (-2%)	Epissage normal et saut de l'exon 6
c.1167A>G	7.11 Cryptique c.1168 non détecté > 8.82	87.89 Cryptique c.1168 non détecté > 89.96	Epissage normal et utilisation du site cryptique c.1168
c.1538-18G>A	6.83 > 5.69 (-16.7%) Cryptique c.1574: 4.71	76.61 > 76.61 (0%) Cryptique c.1574: 85.97	Utilisation du site cryptique c.1574
c.1538-3T>G	6.83 > non détecté Cryptique c.1574: 4.71	76.61 > 74 (-3.4%) Cryptique c.1574: 85.97	Utilisation du site cryptique c.1574
c.1569G>T	6.83 > 6.83 (0%) Cryptique c.1574: 4.71 > 7.22 (+53.4%)	76.61 > 76.61 (0%) Cryptique c.1574: 85.97 > 89.6 (+4.2%)	Utilisation du site cryptique c.1574

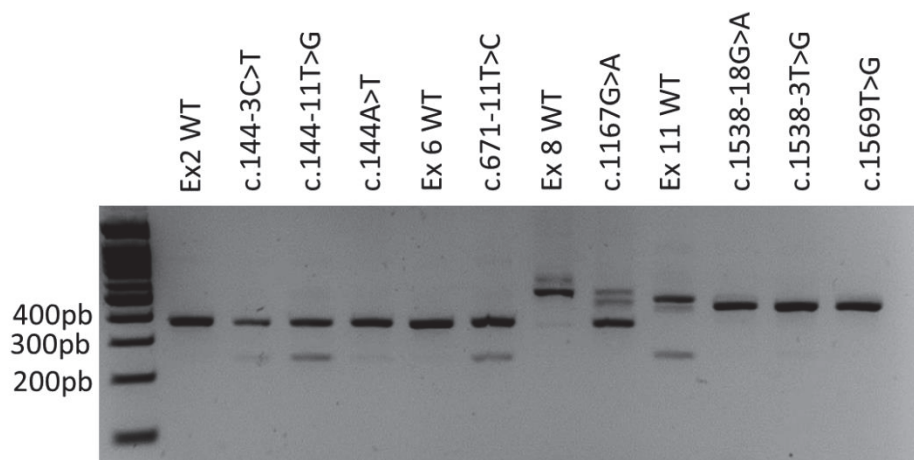
**Tableau 31 : Résultats des analyses *in silico* et *in vitro* obtenus pour les variants situés au niveau d'un site accepteur d'épissage.** Pour l'analyse *in silico*, un variant a été considéré comme affectant l'épissage si les prédictions indiquaient une différence d'au moins 15% pour MaxEntScan et d'au moins 5% pour HSF entre la force du site physiologique et celle du site muté. Dans le cas des sites *de novo* d'épissage, le variant a été considéré comme délétère si la force du site *de novo* était égale ou supérieure à celle du site naturel d'épissage. Pour l'analyse *in vitro*, un variant a été considéré comme délétère si un épissage aberrant était observé dans les deux modèles cellulaires. Les variants grisés correspondent à des discordances entre l'analyse *in silico* et *in vitro*.



Pour la substitution exonique c.1167A>G, l'étude *in silico* était en faveur de la création d'un site accepteur d'épissage *de novo* à la position c.1168 de force supérieure à la force du site naturel (résultat MaxEntScan : score du site *de novo* 8,82 vs 7,11 pour le site naturel). L'étude par minigène de cette mutation confirme la reconnaissance par le spliceosome du site *de novo* en c.1168 entraînant la co-existence de deux transcrits, l'un de 508 pb correspondant à l'épissage correct de l'exon 8 et l'autre de 350 pb correspondant à l'utilisation du site *de novo* en c.1168 (perte des 158 premières paires de bases de l'exon 8) et entraînant l'apparition d'un codon stop (p.Gly338Serfs\*9) (Figure 37 et Figure 38).



**Figure 37 : Résultats de l'analyse *in vitro* effectuée après transfection de cellules HeLa pour les 8 variations situées au niveau d'un site accepteur d'épissage.** Un saut d'exon correspond à une bande de 246 pb dans notre modèle de minigène. WT : Wild Type.



**Figure 38 : Résultats de l'analyse *in vitro* effectuée après transfection de cellules Huh7 pour les 8 variations situées au niveau d'un site accepteur d'épissage.** Un saut d'exon correspond à une bande de 246 pb dans notre modèle de minigène. WT : Wild Type.



L'analyse *in silico* de la dernière variation exonique c.1569G>T montrait une augmentation la force d'un site accepteur cryptique en c.1574 dont le score devenait supérieur à celui du site accepteur naturel de l'exon 11 (résultat MaxEntScan : score du site cryptique 7,22 vs 6,83 pour le site naturel). L'étude par minigène a montré que la mutation c.1569G>T était associée à la présence d'un transcrit unique de 415 pb correspondant à l'utilisation par le spliceosome du site cryptique c.1574 (perte des 36 premiers nucléotides de l'exon 11, p.Val514\_Gly525del13) (Figure 37 et Figure 38).

Les 5 variations introniques touchant un site accepteur d'épissage du gène *F8* sont situées entre les positions -18 et -3 (c.144-3C>T, c.144-11T>G, c.671-11T>C, c.1538-18G>A et c.1538-3T>G). L'analyse *in silico* n'était pas en faveur d'un effet délétère de ces variants. En effet aucun impact sur la force des sites accepteurs naturels n'était prédit pour les variants c.144-3C>T et c.671-11T>C et seulement un des deux algorithmes utilisés prédisait un effet délétère sur l'épissage pour les substitutions c.144-11T>G, c.1538-18G>A et c.1538-3T>G.

L'étude *in vitro* des variants c.144-11T>G et c.144-3C>T a montré la coexistence du transcrit normal (majoritaire) avec un transcrit plus court (246 pb) correspondant à un saut de l'exon 2. Un saut partiel de l'exon 4 a aussi été mis en évidence pour la mutation c.671-11T>C (conséquence protéique : p.Lys225\_Gly263del).

Enfin, les études fonctionnelles sur l'épissage de l'exon 11 des variations c.1538-18G>A et c.1538-3T>G ont montré un résultat similaire à la mutation c.1569G>T : un transcrit unique de 415 pb a été observé. Ce transcrit correspond à l'utilisation par le spliceosome du site accepteur cryptique en c.1574 (perte des 36 premiers nucléotides de l'exon 11, p.Val514\_Gly525del13) (Figure 37 et Figure 38).

## 5. Discussion

La réaction d'épissage est contrôlée par un complexe macromoléculaire appelé spliceosome. Ce complexe est constitué de 5 snRNP contenant un snRNA (snRNA U1, U2, U4, U5 et U6) et de nombreuses autres protéines n'appartenant pas à la famille des snRNP. Le spliceosome reconnaît les sites d'épissage et catalyse les réactions d'excision des introns et de jonction des exons (226). Des

variations génétiques peuvent modifier les signaux d'épissage et donc entraîner l'apparition d'épissage aberrant. Classiquement, l'étude de l'impact d'une variation nucléotidique sur l'épissage associe une étude *in silico* bio-informatique et l'étude des transcrits du gène par RT-PCR et séquençage. Du fait de l'impossibilité d'obtenir des ARNm du gène *F8* issus des tissus producteur de FVIII chez les patients hémophiles A, les études des ARNm du gène *F8* sont réalisées sur les transcrits ectopiques leucocytaires (124,125,205,208–220). Cette approche nécessite de réaliser le prélèvement sanguin à l'aide d'un tube contenant des inhibiteurs de RNase et des stabilisateurs d'ARNm (exemple : tube PAXGENE). Ces tubes étant rarement disponibles dans les services cliniques, des études des transcrits du gène *F8* ont été publiées pour un nombre limité de mutations. Dans l'objectif d'élargir le nombre de patients hémophiles A ayant bénéficié d'une étude de l'impact de leur variation nucléotidique sur l'épissage du gène *F8*, nous avons développé une technique minigène pour confirmer la pathogénicité des variations mises en évidence chez ces patients.

Il est admis qu'aucun logiciel de prédiction bio-informatique n'a une sensibilité ou une spécificité de 100% pour la prédiction de l'impact d'une variation sur l'épissage, à l'exception des mutations affectant les dinucléotides consensus AG/GT. Houdayer *et al.* ont comparé de nombreuses prédictions *in silico* à l'analyse *in vitro* des ARNm pour 272 variations d'épissage (variations situées en dehors des dinucléotides AG/GT) affectant les gènes *BRCA1* et *BRCA2* (227). Ils ont montré que l'utilisation de la combinaison de MaxEntScan avec un seuil de 15% et de HSF avec un seuil de 5% permettait d'obtenir une sensibilité de 96% et une spécificité de 83% pour la prédiction d'un impact d'un variant sur l'épissage. Bien que ces seuils aient été déterminés sur un nombre limité de variations issues de 2 gènes seulement, nous avons décidé d'utiliser ces recommandations pour l'interprétation de nos résultats. Dans ce travail, une bonne corrélation a été observée entre les résultats *in silico* et *in vitro* pour les variations affectant un site donneur d'épissage (89%). A l'inverse, une corrélation très faible a été observée dans le cas des sites accepteurs d'épissage (25%). La grande dégénération de la séquence consensus (à l'exception du dinucléotide AG) des sites accepteurs d'épissage pourrait expliquer la faible sensibilité des logiciels pour la prédiction des impacts de variations localisées près d'un site 3'ss. Le manque de sensibilité de ces logiciels a aussi été souligné par Zimmermann *et al.* dans le cas

des mutations silencieuses, des insertions, des délétions et des substitutions introniques profondes (205). Ainsi, les prédictions *in silico* sont très insuffisantes pour conclure sur la pathogénicité d'un variant d'épissage et une étude fonctionnelle de l'épissage a été réalisée pour tous les variants inclus dans ce travail quels qu'aient été les résultats de l'analyse *in silico*.

L'impact sur l'épissage de tous les variants inclus dans cette étude a été étudié *in vitro* par technique minigène dans deux lignées cellulaires d'origine humaine différentes (HeLa et Huh7). Un variant a été considéré comme ayant un impact sur l'épissage si le profil d'épissage observé sur gel pour le minigène muté était différents de celui observé pour le minigène contrôle pour les deux modèles cellulaires. Il est important d'interpréter les résultats d'étude *in vitro* de l'épissage par minigène avec prudence. En effet de nombreux paramètres importants pour l'épissage correct d'un exon, tels que la vitesse de transcription et les protéines régulatrices de l'épissage, peuvent être différents dans les modèles cellulaires utilisés *in vitro* et dans les cellules productrices de FVIII (228). Ainsi, les transcrits observés *in vitro* (saut d'exon et/ou activation d'un site cryptique) peuvent être différentes des transcrits présents *in vivo* dans les tissus producteurs de FVIII. De plus, des traces d'épissage aberrant ont été observées pour plusieurs minigènes correspondants à des exons non mutés (exons 1, 8, 16, 19). Aucun épissage alternatif n'ayant été décrit à ce jour pour le gène *F8*, ces transcrits minoritaires correspondent sans doute à des artefacts. Leur existence n'empêche pas l'analyse par minigène de variants situés au niveau de ces 4 exons. En effet l'interprétation des tests *in vitro* étant basée sur la comparaison du profil d'épissage du minigène contrôle avec celui du minigène correspondant au variant d'intérêt, les artefacts observés n'ont que peu d'impact. Enfin notre modèle *in vitro* permet de démontrer la pathogénicité d'un variant d'épissage mais le degré de sévérité de l'hémophilie entraîné par ce variant est difficile à estimer. En effet un épissage résiduel normal a été observé *in vitro* pour les mutations c.388+5G>A et c.[787+5G>A;787+12\_13del] alors que ces deux anomalies sont retrouvées chez des hémophiles A sévères.

Quatre des 26 variants étudiés dans ce travail ont déjà été rapportés dans des travaux ayant étudié leur impact sur l'épissage par analyse des transcrits ectopiques du gène *F8*. Des résultats identiques entre l'étude par minigène réalisée dans ce travail et l'étude par RT-PCT des transcrits du gène *F8* des

patients ont été trouvés pour 3 de ces variants (c.388+5G>A, c.601+5G>A et c.1598-18G>A) (209,211). Pour le quatrième variant (c.222G>A), aucun effet délétère sur l'épissage n'a été trouvé dans notre étude. Ce variant a été décrit dans deux études dans lesquelles des analyses par RT-PCR des ARNm du gène *F8* ont été réalisées chez des patients porteurs de cette anomalie. Ces études donnent des résultats contradictoires : aucun impact sur l'épissage n'a été trouvé par Zimmermann *et al.* (205) alors qu'un saut partiel de l'exon 2 du gène *F8* a été mis en évidence par Martorell *et al.* (210). La pathogénicité du variant c.222G>A reste donc encore incertaine.

Dans cette étude, un impact délétère sur l'épissage *in vitro* a été mis en évidence pour 8/10 (80%) des variants affectant le nucléotide consensus G aux positions -1 (dernière base de l'exon) et +5 des sites donneurs d'épissage. Ces 8 mutations (c.388+5G>A, c.601+5G>A, c.670G>A, c.[787+5G>A ; 787+12\_787+13del], c.1009G>A, c.1752G>A, c.5373+5G>T et c.5586G>A) sont donc clairement responsables des phénotypes d'hémophilie A observés chez les patients. Les G aux positions -1 et +5 du site donneur sont très conservés (G présent respectivement dans 80% et 78% des cas à la position -1 et +5) et 70% des mutations du site donneur (en dehors des mutations de la séquence consensus GT) affectent ces deux positions (229). La grande conservation du G à ces deux positions est sans doute lié au fait que les nucléotides -1 et +5 sont les plus importants (après le GT en position +1 et +2) pour l'appariement du snRNA U1 : les G des positions -1, +1 et +5 forment 3 liaisons hydrogènes chacun avec le snRNA U1. Les positions -1 et +5 sont aussi impliquées pour l'appariement des snRNA U5 et U6 (183,230). Ainsi, une variation affectant l'une de ces deux positions peut altérer la reconnaissance du site d'épissage par le spliceosome. Du fait de la discordance observée entre les deux lignées cellulaires lors des analyses *in vitro*, il est difficile de conclure sur la pathogénicité des deux variants affectant la position +5 du site donneur de l'exon 13 (c.2113+5G>A et c.2113+5\_2113+8del). Des analyses des transcrits ectopiques du gène *F8* seront nécessaires chez les patients porteurs de ces variations nucléotidiques.

De nombreux polymorphismes, ayant comme allèle A/G, au niveau de la position +3 du site donneur d'épissage ont été décrits (230). Ceci est probablement dû au fait que la séquence consensus des sites donneur d'épissage comprend soit un G soit un A à la position +3 (177). Le Guédard-

Méreuze *et al.* ont montré que la présence du nucléotide consensus G à la position -1 pouvait compenser la présence d'un nucléotide non consensus (C ou T) à la position +3 d'un site donneur (231). Cependant, en dépit de la présence d'un G en position -1 du site donneur de l'exon 18 du gène *F8*, l'analyse *in vitro* par minigène a montré la présence d'un saut partiel de l'exon 18 pour les mutations c.5998+3A>T et c.5998+3A>C. Ces résultats sont cohérents avec les phénotypes mineurs d'HA observés chez les patients porteurs de ces anomalies.

Les nucléotides consensus des positions +3 et +4 sont moins conservés que ceux des positions -1 et +5. Les nucléotides de ces deux positions ont une contribution moins importante (mais non négligeable) à la force des sites donneurs d'épissage car ils forment seulement deux liaisons hydrogènes chacun avec le snRNA U1 (232). L'évaluation de la pathogénicité d'une variation affectant l'une de ces deux positions est donc assez difficile. Dans ce travail, la pathogénicité des deux mutations affectant la position +4 de l'exon 1 du gène *F8* (c.143+4A>G et c.143+3\_144+4insG) a été clairement démontrée *in vitro*.

L'analyse *in vitro* par minigène des variations situées au niveau d'un site accepteur d'épissage a montré la présence d'épissage aberrant dans tous les cas. La mutation c.1538-18G>A est responsable de la création d'un nouveau AG en amont du site accepteur naturel de l'exon 11 du gène *F8*. La région comprise entre BP et le site accepteur est classiquement dépourvue de dinucléotide AG (188). En effet il est classiquement admis que le spliceosome reconnaît le premier AG localisé en 3' du BP (233). Cependant le nouvel AG créé par la mutation c.1538-18G>A n'est pas utilisé par le spliceosome mais cette mutation entraîne l'activation d'un site accepteur d'épissage cryptique localisé 36 nucléotides en aval du site accepteur naturel de l'exon 11. Une explication possible à cette observation est que le nouveau AG créé est trop proche du BP. En effet une distance minimum de 19 pb entre ces deux éléments est nécessaire (188). La mutation c.1538-18G>A pourrait entraîner la répression du site accepteur de l'exon 11 au profit du site cryptique en c.1574 en modifiant soit des éléments régulateurs soit la structure tridimensionnelle du pré-ARNm.

Les variations silencieuses n'ont par définition aucun impact sur la séquence protéique, en revanche, elles peuvent entraîner des anomalies d'épissage du pré-ARNm. Ce travail a permis de classer 5 des 8 variations silencieuses étudiées (c.1167A>G (p.Gln389Gln), c.1569G>T (p.Leu523Leu), c.1752G>A (p.Gln584Gln), c.5586G>A (p.Leu1862Leu) et c.6066C>T (p.Gly2022Gly)) comme des mutations d'épissage. Ces résultats soulignent la nécessité d'étudier l'impact sur l'épissage des variations silencieuses avant de conclure sur leur pathogénicité. De façon similaire, des variations faux-sens peuvent aussi avoir un impact sur l'épissage comme c'est le cas pour les mutations c.670G>A (p.Gly224Arg) et c.1009G>A (p.Asp337Asn).

Deux patients inclus dans cette étude sont porteurs de deux anomalies génétiques différentes. La mutation c.670G>A est associée à la substitution faux-sens c.3674A>G chez une femme à taux bas de FVIII (FVIII:C = 19%) et la mutation c.1752G>A est associée à la substitution c.-51G>A chez un hémophile A mineur. Notre équipe a précédemment démontré que les variants c.3674A>G et c.-51G>A étaient des polymorphismes rares (104,108). Ce travail a permis de caractériser les mutations c.670G>A et c.1752G>A comme responsables des phénotypes cliniques observés chez les patients. La femme porteuse des variations c.670G>A et c.3674A>G est la nièce d'un hémophile A de sévérité inconnue décédé à l'âge de 2 ans et pour lequel ni dosage de FVIII:C ni échantillon d'ADN ne sont disponibles. Le saut total de l'exon 5 observé lors de l'analyse *in vitro* par minigène de la mutation c.670G>A et le décès précoce du propositus suggèrent fortement que la patiente est conductrice d'hémophilie A sévère. La détermination exacte de l'impact de chacune des anomalies détectées chez un patient est essentielle notamment chez les femmes conductrices afin d'améliorer le conseil génétique.

Aucun impact sur l'épissage n'a été observé lors de l'étude par minigène pour 3 variants étudiés (c.222G>A, c.237C>T et c.240C>T). Ces résultats pourraient suggérer que ces substitutions soient des polymorphismes rares. Cependant il est impossible d'exclure un possible manque de sensibilité de notre technique minigène. L'analyse des ARNm des patients devra être réalisée afin de conclure définitivement sur la pathogénicité de ces 3 variants.

En conclusion, la démonstration de la pathogénicité d'un variant d'épissage permettant d'améliorer le conseil génétique donné à la famille. Ce type d'étude devrait être réalisé pour tous les patients porteurs de variations affectant potentiellement l'épissage du gène *F8* pour lesquels aucun échantillon d'ARNm n'est disponible.

## 6. Annexes

### 1. Liste des amorces utilisées pour la création des minigènes

Minigène	Couples d'amorces
F8 exon 1	5'-TCGAGGATCTCCCCGGGcacaatccagtgaggtaaagtcc-3' 5'-GCGGGCGAGGAGCCCCGGGaatgtttcttggggcccagg-3'
F8 exon 2	5'-CTAAACAGCCACATATGggctgatcagtgatggttattg-3' 5'-CCCCCTCGACCATATGcatctaagtgaggtaagcagttgg-3'
F8 exon 3	5'-CTAAACAGCCACATATGatcttgccctcagagcaacca-3' 5'-CCCCCTCGACCATATGgactggtgtccactacacc-3'
F8 exon 4	5'-CTAAACAGCCACATATGcctgggagcgggagc-3' 5'-CCCCCTCGACCATATGctcagctcctctgatccc-3'
F8 exon 5	5'-CTAAACAGCCACATATGagtggttctcaaatgtaaaa-3' 5'-CCCCCTCGACCATATGatagctgcccagtggaactgag-3'
F8 exon 6	5'-CTAAACAGCCACATATGtgtcatggacaaatggaggg-3' 5'-CCCCCTCGACCATATGtaagatccccagggccagag-3'
F8 exon 7	5'-CTAAACAGCCACATATGgaatggtcaacagtgcaataaag-3' 5'-CCCCCTCGACCATATGcatgagtctactaataactcaagccc-3'
F8 exon 8	5'-CTAAACAGCCACATATgcttggaaagcctaataagagaaaag-3' 5'-CCCCCTCGACCATATGcagcttatatgattcaatcaaggaaggag-3'
F8 exon 11	5'-CTAAACAGCCACATATGaggtgagactttgacttccact-3' 5'-CCCCCTCGACCATATGatccagcaggcacgtttactac-3'
F8 exon 13	5'-CTAAACAGCCACATATGtggtttatgactgtctcctcaca-3' 5'-CCCCCTCGACCATATGtccctgtacctcaaggaagaa-3'
F8 exon 15	5'-CTAAACAGCCACATATGtgcaaaatgcttctcaggca-3' 5'-CCCCCTCGACCATATGacgccaggagttgatatgtaagt-3'
F8 exon 16	5'-CTAAACAGCCACATATGtcccttgcaactttgagttcatacc-3' 5'-CCCCCTCGACCATATGgccagtcacaacctgacct-3'
F8 exon 18	5'-CTAAACAGCCACATATGtatatctgtgggagtggaatc-3' 5'-CCCCCTCGACCATATGatatctgaaattatctgttcatttgc-3'
F8 exon 19	5'-CTAAACAGCCACATATGcctgtttctcctcctctgggg-3' 5'-CCCCCTCGACCATATGctggagatcttcagctttaccaag-3'

Les séquences en majuscule correspondent aux queues ajoutées en 5' des amorces nécessaires à la réaction in-fusion.

### 2. Amorces utilisées pour les réactions de mutagenèse dirigée

Variations	Couples d'amorces
c.144A<T (p.Arg48Ser)	5'-CTTTTACATTTTCAGTTTTCTCCTAGAGTGCC-3' 5'-CTAGGAGGAAAACCTGAAATGTAAGAAGATACG-3'
c.670G>A (p.Gly224Arg)	5'-GCTGTATTTGATGAAAGGAAAAGTTGGC -3' 5'-CTGAGTGCCAACTTTTCTTTTCATCAAATAC -3'
c.1009G>A (p.Asp337Asn)	5'-CTCTTCCCACCAACATAATGGCATGGAAGCTTATG-3' 5'-GCTTCCATGCCATTATGTTGGTGGGAAGAGATATG-3'



3. Liste des amorces utilisées pour le séquençage du plasmide pcDNA3F8<sub>1+13</sub>

Nom du primer	séquence
F8S	5'-CGAGACCCAAGCTGGGTAG-3'
408A	5'-TTCTTTCTCCCTTTGACTGGTC-3'
302S	5'-ATACAGTGGTCATTACACTT-3'
665A	5'-GAGTGCCAACCTTTCCCTTCAT-3'
613S	5'-AAGGAAAAGACACAGACCTT-3'
990A	5'-TCATGTTGGTGGGAAGAGATA-3'
938S	5'-TTACTGCTCAAACACTCTTG-3'
1290A	5'-GCTGAGGGCCATTGTTCAA-3'
1234S	5'-TGGGACTATGCTCCCTTA-3'
1602A	5'-AGTTGGCCCATCTTCTACAGTCACTGTC-3'
1549S	5'-TTGAAGGATTTTCCAATTCT-3'
1900A	5'-AACATAGCCATTGATGCTGT-3'
1785S	5'-TTCTGTATTTGATGAGAACC-3'
5542A	5'-AGGCTTTGCAGACAAACTC-3'
5470S	5'-AACTTTGTCAAGCCTAATGAA-3'
5801A	5'-CATGGAAGCGATAAT-3'
5735S	5'-AAAATATGGAAAGAACTGCA-3'
6105A	5'-GTCTGACACTTATTGCTGTAC-3'
6056S	5'-GCCTTATTGGCGAGACATCTA-3'
6350A	5'-ACATGATGATAAACTGAGAGA-3'
6296S	5'-TTATTCACGGCATCAAGACCC-3'
6549A	5'-AAATCACAGCCCATCAACTCC-3'
6482S	5'-CTCCAATTATTGCTCGATACAT-3'
F8R	5'-GGCACGGGGGAGGGGCAAACAACA-3'

**Section 4 : Description de nouveaux mécanismes moléculaires responsables  
de pathologies hémorragiques héréditaires.**

**Partie 1 : Exonisation d'une séquence *AluY* liée à une anomalie de fixation de l'hnRNP C: un mécanisme original responsable d'hémophilie A mineure**

**Recurrent *F8* intronic deletion in mild haemophilia A causes *Alu* exonization due to the decrease of hnRNP C binding.**

**Y. Jourdy** \*†, A. Janin‡, M. Fretigny \*, C. Negrier \*†¶, D. Bozon ‡, C. Vinciguerra \*†  
Soumis à : American Journal of Human Genetic (en révision)

\* Hospices Civils de Lyon, Hôpital Edouard Herriot, Service d'hématologie Biologique, Lyon, France

† EAM 4174 Hémostase, Inflammation et Sepsis, Université Claude Bernard Lyon 1, France.

‡ Hospices Civils de Lyon, Groupe Hospitalier Est, Laboratoire de Cardiogénétique, Bron, France.

¶ Hospices Civils de Lyon, Hôpital Cardiologique Louis Pradel, Unité d'Hémostase Clinique, Bron, France.

**Keywords:** Haemophilia A, *Alu* elements, hnRNP C, deep intronic mutation

## 1. Résumé

**Contexte :** L'incorporation de séquences introniques dans les ARNm mature est un mécanisme responsable de maladies génétique. De nombreux pseudo-exons responsables de différentes maladies génétiques contiennent des séquences répétées telles que des séquences *Alu*.

**Objectif :** L'objectif de cette étude est la description d'un mécanisme moléculaire original par lequel une délétion intronique entraîne l'exonisation d'une séquence *Alu*.

**Matériels et méthodes :** Des analyses de transcrits ectopiques leucocytaires du gène *F8* ont été réalisées chez deux patients hémophiles A mineurs chez lesquels aucune anomalie génétique n'était connue à ce jour. Des modèles cellulaires (HeLa et Huh7) *in vitro* de type minigène ont été utilisés pour étudier l'impact de la délétion c.2113+461\_2113+473del sur l'exonisation d'une séquence *AluY*.

**Résultats :** la délétion intronique profonde (c.2113+461\_2113+473del) a été identifiée chez les deux propositus. Les études *in vivo* et *in vitro* des transcrits ont montré la co-existence chez les patients de deux transcrits, l'un normal et l'autre plus long. Dans ce dernier transcrit, une insertion de 122 bp (c.2113\_2114ins2113+477\_2113+598) entre les exons 13 et 14 a été mise en évidence. Cette insertion entraîne un décalage du cadre de lecture (p.Gly705Aspfs\*37). L'étude de la séquence de ce pseudoexon a montré qu'il correspondait à une séquence *AluY* antisens et que la délétion était responsable de la perte de 13 des 25 T formant la queue poly-(T) de cette séquence *AluY*. L'hnRNP C se fixe sur la queue poly-(T) des séquences *Alu* inversées du génome afin d'éviter leur exonisation. L'introduction de mutation par mutagenèse dirigée au niveau de la queue poly-(T) de cette séquence *AluY* ou l'inhibition de la production d'hnRNP C à l'aide d'un siRNA permet de reproduire *in vitro* l'effet de la délétion c.2113+461\_2113+473del. Cette délétion a été recherchée dans les 114 familles d'HA mineurs de notre cohorte pour lesquelles aucune anomalie génétique n'était connue. La présence d'une délétion dans la queue poly-(T) de cette séquence *AluY* a été trouvée dans plus de la moitié des cas (n = 61/114 ; 54%).

**Conclusion :** Cette étude décrit une nouvelle délétion localisée dans l'intron 13 du gène *F8*. Cette anomalie génétique est responsable de l'exonisation d'une séquence *AluY* et est retrouvée dans 6,1% des HA mineures de notre cohorte.

## 2. Rappels bibliographiques : les séquences *Alu*.

Les séquences *Alu* sont des séquences d'ADN d'environ 300 pb appartenant à la famille des petits éléments nucléaires intercalés (SINE, short interspersed nuclear elements) caractérisées par la présence d'un site de restriction pour l'enzyme *AluI*.

### 1. Les séquences *Alu* au sein des éléments transposables

Les éléments transposables sont des séquences d'ADN capables de se déplacer et de se multiplier dans le génome. Ils constituent environ 45% du génome humain (234). Les éléments transposables sont classés en transposons à ARN (classe I) ou à ADN (classe II) en fonction de la nature de la molécule intermédiaire nécessaire à leur mécanisme d'amplification.

Les transposons de classe I, aussi appelés rétrotransposons, se répliquent à l'aide d'un processus appelé rétrotransposition. Lors de ce processus, le rétrotransposon est transcrit en un ARN intermédiaire qui sera, après rétrotranscription en ADN, inséré à une autre localisation dans le génome (235). Les rétrotransposons sont subdivisés en deux familles en fonction de la présence ou non de séquences LTRs (Long Terminal Repeats) à leurs extrémités 5' et 3'. Les séquences LTRs jouent un rôle dans la transcription et l'intégration dans le génome des rétrotransposons de type LTR. A ce jour, seuls certains rétrotransposons de type non-LTR semblent encore actifs chez l'Homme. Les rétrotransposons de type non-LTR peuvent être eux-mêmes divisés en éléments autonomes et non-autonomes en fonction de leur capacité à coder les enzymes et les protéines nécessaires à leur rétrotransposition. Les séquences LINEs (long interspersed elements) appartiennent au groupe des rétrotranspositions autonomes. Les séquences LINEs possèdent deux cadres de lecture et codent deux protéines nécessaires à leur cycle d'amplification: ORF1p, une protéine chaperonne ayant la capacité de se lier aux ARN, et ORF2, une protéine possédant une activité endonucléase et transcriptase inverse.

A l'inverse des éléments autonomes, les rétrotransposons de type non-LTR et non-autonomes doivent recruter les facteurs nécessaires à leur amplification et à leur intégration dans le génome. Chez l'Homme, ce sont classiquement les protéines ORF1p et ORF2p codées par les éléments LINE-1 qui sont utilisées par ce type de rétrotransposon (236). Les séquences *Alu* appartiennent au groupe des

rétrotransposons non-autonomes. Les séquences *Alu* constituent l'élément transposable le plus répandu dans le génome humain avec environ 1 million de copies (correspondant à 11 % du génome humain) (237).

## 2. Structure des séquences *Alu*

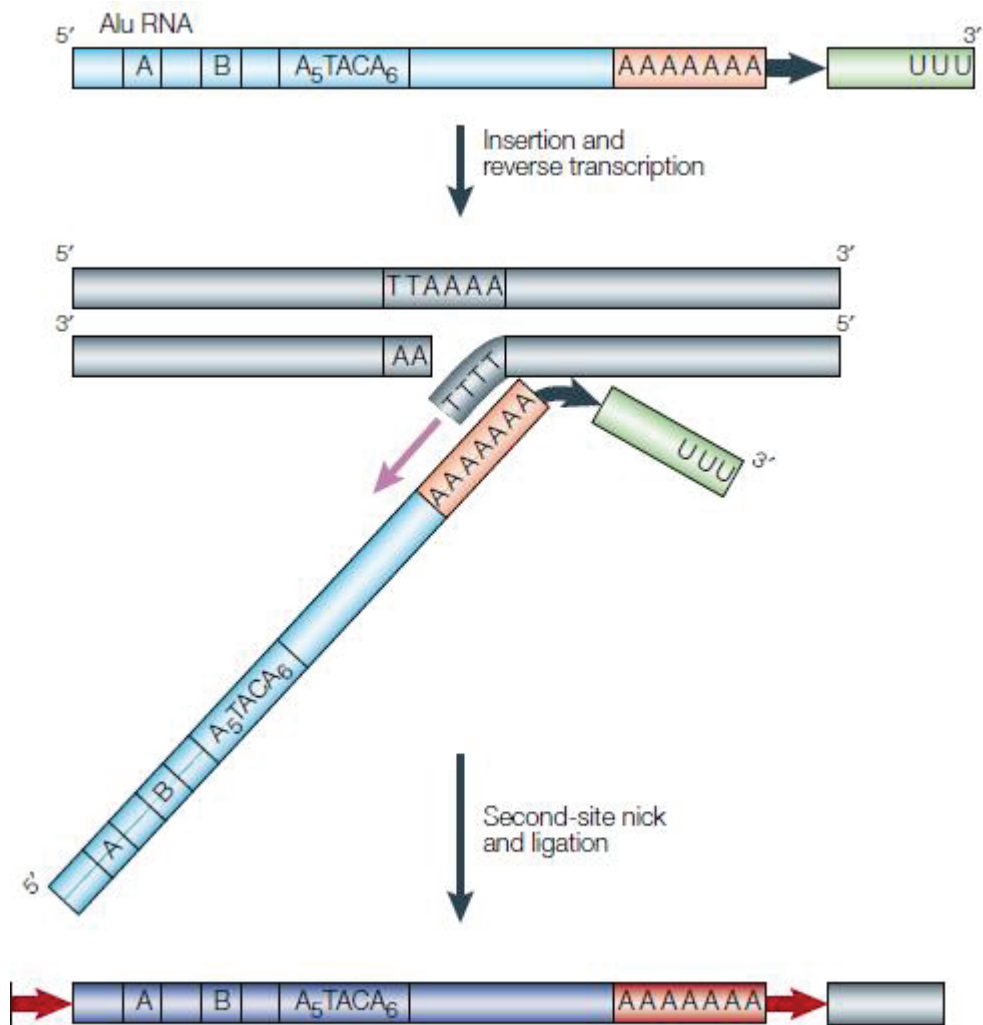
Les séquences *Alu* dérivent d'un gène ancestral codant l'ARN 7SL (238). Ce sont des éléments d'environ 300 pb ayant une structure dimérique. Les deux monomères (droit et gauche) composant une séquence *Alu* présentent une grande homologie de séquence. Ils sont séparés par une région riche en A. Les séquences *Alu* possèdent une séquence promotrice pour l'ARN polymérase III (boîtes A et B) dans leur région 5' et une queue poly-(A) de taille variable (20 à 100 adénines) à leur extrémité 3'.



**Figure 39 : structure d'une séquence *Alu* (239).** Les séquences *Alu* sont formées de deux monomères ayant une grande homologie de séquence, séparés par une courte région riche en adénines (AA). Elles se terminent par une queue poly-(A). Les séquences *Alu* possèdent une région promotrice pour l'ARN polymérase III (boîtes A et B) mais n'ont pas de signal de fin de transcription. Pour terminer la transcription, l'ARN polymérase III utilise des régions composées de répétitions de T situées à des distances variables en aval des séquences *Alu*.

## 3. Mécanisme d'amplification et d'intégration des séquences *Alu*

Les séquences *Alu* actives du génome sont transcrites en ARN par l'ARN polymérase III. Les séquences *Alu* ne possèdent pas de séquence de terminaison de la transcription, l'ARN polymérase arrête la transcription dès qu'elle rencontre la séquence TTTT dans le voisinage de la séquence *Alu* (Figure 39). L'étape de rétrotranscription est réalisée au niveau du site d'insertion. L'enzyme ORF2p coupe l'un des brins d'ADN au niveau d'une séquence consensus AA/TTTT, la queue poly-(A) de la séquence *Alu* s'hybride au niveau de la répétition de T de l'ADN génomique. Cette séquence riche en T sert d'amorce pour la réaction de rétrotranscription catalysée par l'enzyme ORF2p. Les mécanismes impliqués pour la ligation de l'extrémité 5' de la nouvelle séquence *Alu* ainsi que l'intégration cette séquence sur le second brin d'ADN sont inconnus à ce jour (239).



**Figure 40 : mécanisme de rétrotransposition des séquences *Alu* (240).** Les séquences *Alu* actives du génome sont transcrites en ARN par l'ARN polymérase III. Les séquences *Alu* ne possèdent pas de séquence de terminaison de la transcription, l'ARN polymérase arrête la transcription dès qu'elle rencontre la séquence TTTT dans le voisinage de la séquence *Alu*. L'étape de rétrotranscription est réalisée au niveau du site d'insertion. L'enzyme ORF2p coupe l'un des brins d'ADN au niveau d'une séquence consensus AA/TTTT, la queue poly-(A) de la séquence *Alu* s'hybride au niveau de la répétition de T de l'ADN génomique. Cette séquence riche en T sert d'amorce pour la réaction de rétrotranscription catalysée par l'enzyme ORF2p. Les mécanismes impliqués pour la ligation de l'extrémité 5' de la nouvelle séquence *Alu* ainsi que l'intégration cette séquence sur le second brin d'ADN sont inconnus à ce jour

Parmi le million de séquences *Alu* du génome humain, seule quelques-unes ont encore la capacité de s'amplifier (241). La capacité d'une séquence *Alu* à se rétrotransposer dépend de nombreux facteurs dont :

- La région du génome où s'est intégrée la séquence *Alu*. Le promoteur de l'ARN polymérase III présent dans ces éléments est insuffisant pour assurer une transcription efficace. Seules les



séquences *Alu* ayant été insérées dans des régions du génome contenant des séquences enhancers sont donc actives (242).

- La méthylation des îlots CpG. Les séquences *Alu* possèdent au minimum 24 dinucléotides CpG dans leur séquence. La méthylation de ces îlots pourrait influencer l'efficacité de la transcription des séquences *Alu* par l'ARN polymérase III.
- Les séquences *Alu* accumulent des mutations au cours du temps, notamment au niveau des dinucléotides CpG (par déamination des méthylcytosines). Ces mutations peuvent altérer la région promotrice de l'ARN polymérase III, la structure de l'ARN ou encore les interactions entre l'ARN et les protéines cellulaires intervenant dans le phénomène de rétrotransposition (243).
- La taille de la queue poly-(A) est un facteur important pour la rétrotransposition. La réduction de cette taille par accumulation de mutations est un mécanisme responsable de l'inhibition de séquences *Alu* (241).
- La taille de l'extrémité 3' de l'ARN des séquences *Alu* influence la capacité d'amplification des séquences *Alu* : les extrémités 3' de grande taille sont un facteur délétère (241).

L'accumulation de tous ces éléments a contribué à la perte d'activité de la plupart des séquences *Alu* du génome humain au cours de l'évolution.

#### 4. Multiplification et diversité des séquences *Alu*

Depuis leur apparition dans le génome des primates il y a environ 65 millions d'années, les séquences *Alu* se sont énormément multipliées et diversifiées. En moyenne, depuis l'apparition des séquences *Alu*, il est estimé qu'une nouvelle insertion de séquence *Alu* a eu lieu toutes les 21 naissances. Cependant, la vitesse de multiplication des séquences *Alu* dans le génome a varié au cours du temps et la majorité des séquences *Alu* du génome humain était déjà présent il y a 40 millions d'années. Pendant la période où la multiplication de ces rétrotransposon était la plus rapide, il est estimé qu'une nouvelle insertion de séquence *Alu* avait lieu à chaque naissance alors qu'actuellement cette moyenne est estimée à environ une insertion toutes les 200 naissances (244,245).

En fonction de leur séquence nucléotidique, les séquences *Alu* sont classées en différentes sous-familles (plus de 200 actuellement) dont les principales sont *AluJ*, *AluS* et *AluY* (246). Parmi ces trois lignées principales, la branche *AluJ* est la plus ancienne et la branche *AluY* est la plus récente. Les sous-familles de séquences *Alu* sont apparues au cours de l'évolution à la suite de l'apparition de mutations ponctuelles sur une séquence *Alu* ancestrale qui avait une grande capacité de rétrotransposition.

Il semblerait qu'au cours de l'évolution, les sous-familles de séquences *Alu* très actives, donc potentiellement délétères, ont été rapidement sélectionnées négativement en accumulant des mutations délétères à leur activité. Les lignées *Alu* ayant une activité très faible ont seules persisté sur de très longues périodes. La limitation de leur capacité à se multiplier est une stratégie évolutive commune à plusieurs familles de rétrotransposons (244). De manière intéressante, certaines lignées *Alu* s'étant maintenues en activité dans le génome pendant de très longues périodes grâce à cette stratégie ont conservé leur capacité à générer à un instant donné un grand nombre de nouvelles copies. Ainsi la vitesse d'amplification de certaines sous-familles a varié au court du temps. C'est le cas par exemple des séquences *AluYb*. Elles sont apparues il y a 18-25 millions d'années chez les Hominoïdes (Homme et grands singes). De manière surprenante, ces séquences se sont multipliées ces quelques derniers millions d'années spécifiquement dans le génome humain (apparition d'environ 2000 copies) alors que sur la même période, seule une poignée de séquences *AluYb* sont apparues dans le génome des grands singes. Ainsi les séquences *AluYb* ont persisté dans le génome des Hominoïdes avec une activité très faible voire nulle pendant 15-20 millions d'années puis leur activité a repris spécifiquement chez l'Homme, entraînant la formation d'un grand nombre de copies (244).

##### 5. Rôle des séquences *Alu* dans l'évolution des génomes

Les séquences *Alu* ont joué un rôle très important dans l'évolution du génome des primates. Pendant les 6 derniers millions d'années, 5530 séquences *Alu* ont été nouvellement insérées dans le génome humain (247). La plupart de ces insertions ont été la conséquence d'une rétrotransposition classique. Dans de rares cas, des séquences *Alu* a été insérées dans le génome à la suite de

recombinaisons homologues. Ces insertions *de novo* de séquences *Alu* ont participé à l'évolution du génome humain par différents mécanismes :

- Interruption de gène lors d'une insertion dans une région exonique.
- Création d'épissage alternatif : les séquences *Alu* insérées dans des régions introniques peuvent donner naissance à des épissages alternatifs. En effet, elles contiennent de nombreux sites cryptiques d'épissage qui peuvent être reconnus par le splicéosome. L'exonisation de séquences *Alu* et/ou la rétention de séquences introniques dans les transcrits matures de gène a participé à la complexification du protéome humain (248).
- Modification de l'expression des gènes par édition d'ARNm. L'édition des ARNm est le processus de modification post-transcriptionnel des ARNm le plus courant. Lors de ce processus, une enzyme appelée Adenosine Deaminase Acting on RNA (ADAR) catalyse la désamination des adénosines en inosines sur les séquences d'ARNm double brin. L'inosine est interprétée comme une guanosine par le splicéosome ou lors de la traduction en protéine. Les séquences *Alu* localisées dans les introns d'un gène favorisent de par leur structure dimérique la création de boucles ou d'épingles à cheveux permettant l'action de l'ADAR. De la même manière, deux séquences *Alu* proches l'une de l'autre peuvent s'apparier et former une structure en épingle à cheveux. L'édition des ARNm peut entraîner l'apparition d'un nouvel épissage alternatif (249,250).
- Modification du signal de polyadénylation. Les séquences *Alu* peuvent réguler la fonction des gènes en introduisant dans ces derniers de nouveaux signaux de polyadénylation. En effet, ces rétrotransposons contiennent au moins deux séquences signal (AATAAA) potentielles de polyadénylation. Ainsi, parmi les 10 000 séquences *Alu* présentes dans les régions 3' non traduites des gènes, 107 contiennent un site actif de polyadénylation (251).

Les séquences *Alu* sont aussi impliquées dans certains grands réarrangements au cours de l'évolution du génome humain. On dénombre ainsi 515 délétions liées à des séquences *Alu* chez l'Homme. Parmi ces délétions, 24 sont liées à l'insertion d'une nouvelle séquence *Alu* et 492 font suite à une recombinaison impliquant des séquences *Alu*. Au total, ce sont environ 416 kb qui ont été

délétés dans le génome humain depuis la divergence entre les Hommes et les Chimpanzés il y a environ 6 millions d'années (234). Les séquences *Alu* sont fréquemment impliquées dans les phénomènes de recombinaison du fait de deux caractéristiques : leur grande homologie de séquence et le nombre important de copies. Les séquences *AluY*, plus récentes et donc ayant accumulées moins de mutations que les lignées plus anciennes (*AluJ*), présentent une plus grande homologie de séquence et sont donc les plus fréquemment impliquées dans des phénomènes de recombinaison (252).

#### 6. Implication des séquences *Alu* en pathologie humaine

De la même manière qu'elles ont participé à la modification du génome humain au cours du temps, les séquences *Alu* sont aussi la cause de nombreuses maladies génétiques. Ainsi, il est estimé que l'insertion de nouvelles séquences *Alu* est responsable de 0,1% des maladies génétiques humaines et que les réarrangements génomiques liés aux séquences *Alu* comptent pour 0,3% (234). A ce jour, une centaine de cas de maladies génétiques causées par des séquences *Alu* a été rapporté dans la littérature (234,253). Dans le cadre des pathologies hémorragiques, 4 insertions *de novo* de séquence *Alu* au niveau des régions codantes des gènes *F8* et *F9*, entraînant un décalage du cadre de lecture, ont été rapportées chez 2 hémophiles A et 3 hémophiles B sévères (141,254–257). Une insertion d'une séquence *AluYb9* dans l'intron 18 du gène *F8* entraînant le saut complet de l'exon 19 a aussi été décrite chez un hémophile A sévère (218).

Les séquences *Alu* peuvent aussi être responsables de maladie génétique par création d'exon cryptique. Ainsi une dizaine de cas d'exonisation de séquence *Alu* a été rapporté dans la littérature (258). Dans la majorité des cas, des mutations introniques profondes créent un site *de novo* d'épissage cryptique et entraînent l'activation d'un des sites cryptiques d'épissage présent dans la séquence *Alu*. Quarante-vingt-dix pourcents des séquences *Alu* impliquées dans la création de pseudoexon en pathologie humaine sont orientées en antisens (258). La présence d'une queue poly pyrimidine poly-T (séquence complémentaire de la queue poly-A) suivie d'un trinuécléotide CAG mimant les motifs caractéristiques d'un site accepteur d'épissage facilite probablement l'exonisation des séquences *Alu* inversées (253).

### 3. Introduction

Les principaux types de mutations du gène *F8* trouvées chez les hémophiles A mineurs sont des mutations faux-sens (75%) suivies par les mutations des sites d'épissages (9%). Cependant dans environ 10% des cas d'hémophilies A mineures, aucune mutation n'est mise en évidence (105,259). Des analyses des ARNm ectopiques leucocytaires du gène *F8* ont été réalisées chez plusieurs de ces patients et ont permis de montrer dans certains cas la présence de séquences introniques dans les ARNm mature du gène *F8* (123,213).

L'incorporation de séquence intronique dans les ARNm mature est une cause méconnue responsable de maladie génétique. Le plus souvent cet évènement moléculaire est causé par une mutation intronique profonde qui crée un site d'épissage donneur ou accepteur *de novo* entraînant l'activation d'un exon cryptique (pseudoxon) qui sera inclus dans l'ARNm mature par le spliceosome (260,261). D'autres mécanismes possibles à l'origine d'activation d'exon cryptique ont été décrits : l'élimination ou la création de séquences régulatrices (262) ou la création d'un site de branchement par une mutation intronique profonde (263). Plusieurs pseudoxons décrits en pathologie humaine contiennent des séquences répétées telles que des séquences *Alu* (258).

Les séquences *Alu* sont la famille de rétrotransposon la plus répandue avec environ 1 million de copie répartie dans tout le génome humain (237). Les séquences *Alu* sont des éléments d'environ 300 pb composés de deux domaines homologues séparés par une courte région riche en adénines. L'extrémité 3' des séquences *Alu* est constituée d'une longue queue poly-A composée d'une répétition d'environ 20 à 100 adénines (253). Les séquences *Alu* ne sont pas réparties de façon uniforme dans le génome, elles sont préférentiellement localisées dans les régions riches en gènes et notamment au niveau des introns des gènes. Dans les introns, ces séquences peuvent être orientées soit dans le même sens soit en sens inverse par rapport au sens de transcription du gène (264). En 2010, une étude a montré que 90% des séquences *Alu* impliquées dans la création de pseudoxon en pathologie humaine étaient orientées en sens inverse (258). La présence d'une queue poly pyrimidine poly-T (séquence complémentaire de la queue poly-A) suivie d'un trinuécléotide CAG mimant les motifs caractéristiques d'un site accepteur d'épissage facilite probablement l'exonisation des

séquences *Alu* inversées (253). Ainsi, un très grand nombre de séquences *Alu* localisées dans les introns des gènes possèdent des séquences correspondant parfaitement aux définitions des sites accepteurs et donneurs d'épissage (264,265). Cependant il existe des mécanismes de régulation dont l'objectif est de protéger le transcriptome contre l'exonisation aberrante de séquences *Alu*. L'hRNP C (heterogeneous nuclear riboprotein C1/C2) appartient à l'un des plus importants de ces mécanismes. Cette protéine se lie sur les répétitions ininterrompues d'au moins 5 Uridines (telle que la queue poly-T d'une séquence *Alu* inversée au niveau d'un pré-ARNm) et rentre en compétition avec le facteur d'épissage U2AF65 afin de masquer les séquences *Alu* inversées pour empêcher leur reconnaissance par le spliceosome (266–268).

Récemment nous avons étudiés les transcrits ectopiques leucocytaires du gène *F8* chez des hémophiles A chez lesquels aucune anomalie génétique n'avait été trouvée. Nous avons mis en évidence chez deux hémophiles A mineurs non apparentés l'exonisation d'une séquence *Alu* entre les exons 13 et 14 du gène *F8*. L'objectif de ce travail a été de déterminer la responsabilité de cette anomalie génétique et le mécanisme physiopathologique permettant d'expliquer ce phénomène.

#### **4. Matériels et méthodes**

##### **1. Patients**

Les deux cas index sont des hémophiles A mineurs non apparentés chez lesquels aucune mutation n'a été détectée lors du séquençage de la totalité des exons, des jonctions intron/exon, de la région promotrice et de la région 3' non traduite du gène *F8* ainsi que lors de la recherche d'une anomalie du nombre de copies par MLPA (Multiplex ligation probe amplification). Les principaux diagnostics différentiels de l'HA mineure tels que la maladie de willebrand de de type 1 et 2N et le déficit combiné en FV et FVIII ont été exclus. Aucun de ces deux patients n'a développé d'inhibiteur et le patient 1 est parfaitement répondeur au traitement par la desmopressine. Les caractéristiques cliniques et biologiques des patients sont présentées dans le Tableau 32.

Patients	1	2
Age (années)	56	11
FVIII:C (%)	24	40
FVIII:CR (%)	23	41
FVIII:Ag (%)	17	37
VWF:Rco (%)	61	62
VWF:Ag (%)	50	88
VWF:FVIII B	Normal	Normal
FV (%)	81	95
Inhibiteurs*	Neg	Neg
Réponse à la desmopressine	Oui	Inconnue

**Tableau 32 : caractéristiques cliniques et biologiques des deux cas index.** FVIII:C, activité du FVIII mesurée par une technique coagulométrique FVIII; FVIII:CR, activité du FVIII mesurée par une technique chromogénique; FVIII:Ag, FVIII antigène; VWF:RCo, Activité co-facteur de la ristocétine du facteur von Willebrand; VWF:Ag, mesure antigénique du facteur von Willebrand; VWF:FVIII B, mesure de la liaison du facteur VIII au facteur von Willebrand; Neg, négatif.\* Recherche réalisée par la technique Bethesda, un résultat <0,6U/ml a été considéré comme négatif.

Cinquante-cinq sujets témoins (51 femmes et 4 hommes) ainsi que 94 hémophiles A mineurs chez lesquels aucune mutation n'a été trouvée ont aussi été inclus dans cette étude. Toutes les personnes incluses dans ce travail ont signé un consentement éclairé pour la réalisation d'analyses génétiques.

## 2. Analyse des transcrits ectopiques leucocytaires du gène *F8*

### **Prélèvements et extraction des ARNm totaux leucocytaires**

Les prélèvements sanguins ont été réalisés sur tube PAXgene (PAXgene blood RNA tube, PreAnalytix, Hombrechtikon, Suisse). Les ARNm totaux leucocytaires ont été extraits et purifiés à l'aide du kit PAXgene blood RNA kit (PreAnalytix) en suivant les recommandations du fournisseur.

### **Réalisation de la réaction de RT**

Pour la réaction de RT, l'ADNc du gène *F8* a été divisé en 4 régions couvrant la totalité des jonctions intron/exon : région A (exon 1 à 7), région B (exon 7 à 14), région C (exon 14 à 22) et région D (exon 22 à 26). Nous avons utilisé comme contrôle interne de la réaction le gène de la  $\beta$ -actine (*ACTB*). La liste des amorces utilisées pour l'analyse des transcrits ectopiques leucocytaires du gène *F8* est disponible en annexe de cette section. La réaction de RT a été réalisée en utilisant le kit PrimeScript Reverse Transcription (Clontech, Saint-Germain-en-Laye, France). Dans un premier temps, 2 $\mu$ l de chaque amorce antisens (solution à 10pM) spécifique d'une des 4 régions de l'ADNc du gène *F8* (F8AR, F8BR, F8CR et F8DR) et 2 $\mu$ l de l'amorce antisens du gène *ACTB* ( $\beta$ -actineR) ont été mélangés (obtention d'un pool d'amorces à 2pM). Ensuite 1 $\mu$ g d'ARNm totaux a été dénaturé à 65°C

pendant 5 minutes en présence de 1µl du pool d'amorces à 2pM et 1µl de dNTP (2mM chacun). La réaction de RT a été réalisée dans une seconde étape. La composition du mélange réactionnel et les conditions de réaction sont présentées dans les Tableau 33 et Tableau 34 :

Réactif	Volume (µl)
Mélange ARNm/amorces/dNTP	10
5X PrimeScript Buffer	4
RNase Inhibitor (400U/µl)	0,5 (20U)
PrimeScript Reverse Transcriptase (200U/µl)	0,5 (100U)
Eau RNase Free	5

**Tableau 33 : Composition du mélange réactionnel pour la réaction de RT.**

	Température (°C)	Temps (minutes)	Nombre de cycles
Incubation	30	10	
Elongation	50	60	1
Inactivation de la transcriptase inverse	70	15	

**Tableau 34 : Conditions expérimentales pour la réaction de RT.**

L'amplification de l'ADNc du gène *ACTB* a permis de contrôler à la fois la qualité de la réaction de RT et la pureté des ARNm extraits. La composition du mélange réactionnel et les conditions expérimentales sont indiquées dans les Tableau 35 et Tableau 36 :

Réactifs	Volume (µl)
Tampon Qbiogene with MgCl <sub>2</sub> 10X	2,5
β-actineS [10pM]	1,25
β-actineR [10pM]	1,25
dNTP [2,5 mM de chaque dNTP]	1
Taq DNA polymerase MPbiomedicals [5U/µl]	0,25
Eau	17,75
RT-PCR	1

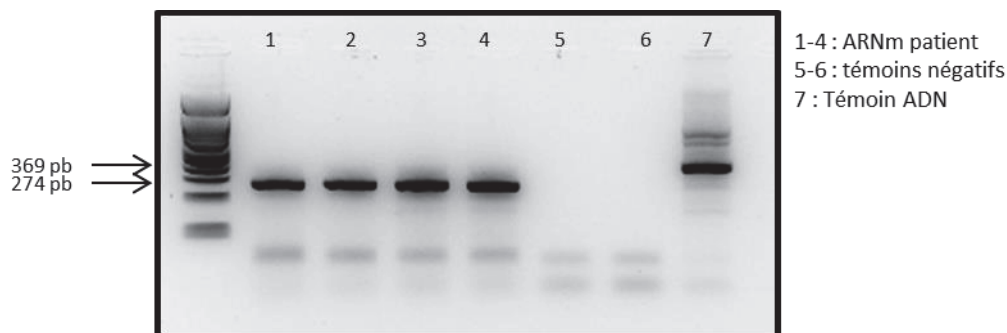
**Tableau 35 : Composition du mélange réactionnel pour l'amplification des transcrits du gène *ACTB* (contrôle interne de la réaction).**

	Température (°C)	Temps (minutes)	Nombre de cycles
Activation	94	5	1
Dénaturation	94	1	
Hybridation	64	1	30
Elongation	72	1	
Elongation finale	72	4	1

**Tableau 36 : Conditions expérimentales pour l'amplification des transcrits du gène *ACTB* (contrôle interne de la réaction).**

Les produits de PCR ont été contrôlés par migration sur gel d'agarose 2% (120 V, 30 minutes). Une bande unique de 274pb correspond au transcrit du gène *ACTB*. Une bande de 369 pb, correspondant au gène *ACTB*, est visible en cas de contamination de l'échantillon par de l'ADN (Figure 41).





**Figure 41 : Exemple d'amplification des transcrits du gène *ACTB*.** Une bande de 274 pb est visible pour les échantillons d'ARNm, ces derniers ne sont pas contaminés par de l'ADN. Une bande de 369 pb est visible pour l'ADN témoin.

### Amplification des produits de RT

Les produits de RT ont été amplifiés par deux étapes de PCR successives en utilisant la technique de la PCR nichée. Dans la première étape, les 4 grandes régions A, B, C, D définies ci-dessus ont été amplifiées à l'aide d'amorces spécifiques puis dans la seconde étape, chaque région a été divisée en deux ou trois fragments (pour un total de 9 fragments). La liste des amorces utilisées est présentée en annexe de cette section. La composition du mélange réactionnel et les conditions de PCR sont présentées les Tableau 37 et Tableau 38 (identique pour les deux étapes).

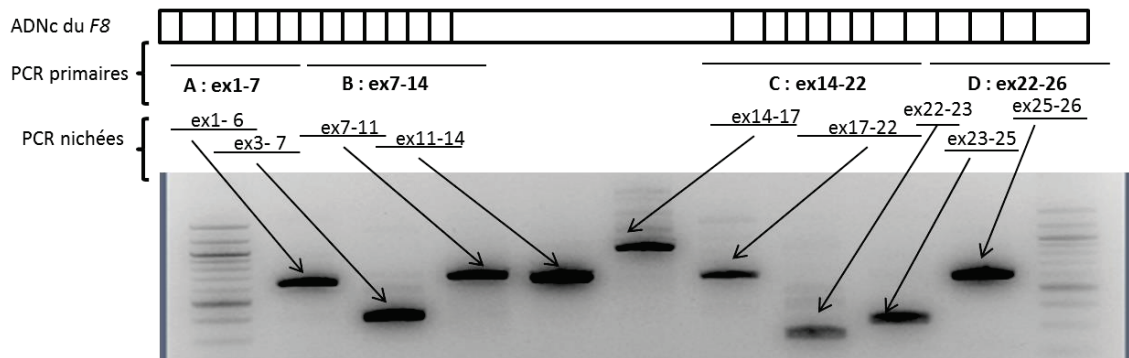
Réactifs	Volume (µl)
Tampon PrimeSTAR GXL 5X	10
Amorce sens [10pM]	1,5
Amorce antisens [10pM]	1,5
dNTP [2,5 mM de chaque dNTP]	4
PrimeSTAR GXL DNA polymerase 1,25U/µl	1
Eau	28
Echantillon (RT-PCR ou 1 <sup>ère</sup> PCR)	4

**Tableau 37 : Composition du mélange réactionnel pour l'amplification des transcrits des produits de RT-PCR.**

	Température (°C)	Temps	Nombre de cycles
Activation	98	5	1
Dénaturation	98	10 s	30
Hybridation	55	15 s	
Elongation	68	2 min	
Elongation finale	68	4	1

**Tableau 38 : Conditions expérimentales pour l'amplification des produits de RT-PCR.**

Les produits de PCR ont été analysés par migration sur gel d'agarose 1% (70 V, 90 minutes) (Figure 42). En cas de profil anormal, les fragments d'intérêt ont été extraits du gel à l'aide du kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, Hoerd, France) et séquencés.



**Figure 42 : Analyse des transcrits du gène *F8* par migration sur gel d'agarose 1%. L'ADNc du gène *F8* est divisé en 4 régions (A à D) chevauchantes. Chaque région a été elle-même divisée en deux (A à C) ou trois (D) segments.**

### 3. Séquencage de la région d'intérêt de l'intron 13 du gène *F8*.

Pour le séquençage de l'intron 13 du gène *F8*, la région comprise entre les positions *F8* c.1904-202 et *F8* c.2113+770 a été amplifiée par PCR en utilisant des amorces spécifiques (annexes de cette section). Les conditions de PCR sont identiques à celles utilisées pour l'étude des transcrits ectopiques du gène *F8* (Tableau 37 et Tableau 38). Le séquençage de la région d'intérêt (*F8*c.1978\_c.2113+675) a ensuite été réalisé à l'aide d'un second couple d'amorces (annexes de cette section).

### 4. Analyse des microsatellites situés dans le gène *F8*.

Six polymorphismes intragéniques (microsatellites) ont été étudiés dans ce travail. Ces marqueurs sont situés dans les introns 1 (STR1), 9 (STR9), 13 (STR13), 22 (STR22) et 25 (STR25A et STR25B) du gène *F8* (269,270). Les descriptions de ces différents marqueurs sont présentées dans le Tableau 39.

Marqueur	Position génomique (hg38)	Type de répétition
STR 1	chrX:155,002,613-155,002,646	(CA) <sub>n</sub>
STR 9	chrX:154,964,082-154,964,109	(CA) <sub>n</sub>
STR 13	chrX:154,935,982-154,936,024	(AC) <sub>n</sub>
STR 22	chrX:154,875,737-154,875,774	(GT) <sub>n</sub>
STR 25 A	chrX:154,850,084-154,850,120	(TG) <sub>n</sub>
STR 25 B	chrX:154,854,591-154,854,619	(TG) <sub>n</sub>

**Tableau 39 : Description des marqueurs intragéniques du gène *F8* utilisés dans ce travail.**

Les réactions de PCR ont été réalisées en utilisant des amorces marquées à la fluorescéine (FAM). La composition des mélanges réactionnels et les conditions de PCR sont indiquées dans les Tableau 40, Tableau 41 et Tableau 42.

Réactifs	Volume (µl)
Tampon Optimase 10X	5
ADN patient (50ng/µl)	1
Amorces sens [25pmol/µl]	1
Amorces antisens [25pmol/µl]	1
dNTP [2,5 mM de chaque dNTP]	4
Optimase Polymerase [2,5U/µl]	0,5
Eau	37,5

**Tableau 40 : Composition du mélange réactionnel pour l'analyse des polymorphismes intragéniques du gène *F8*.**

	Température (°C)	Temps	Nombre de cycles
Dénaturation initiale	95	7 min	
Dénaturation	95	50 s	
Hybridation	62	50 s	10
Elongation	72	50 s	
Dénaturation	95	50 s	
Hybridation	59	50 s	25
Elongation	72	50 s	
Elongation finale	72	10 min	

**Tableau 41 : Conditions expérimentales pour l'amplification par PCR des marqueurs STR1, STR9, STR25A et STR25B.**

	Température (°C)	Temps	Nombre de cycles
Dénaturation initiale	95	7 min	
Dénaturation	95	50 s	
Hybridation	62	50 s	10
Elongation	72	50 s	
Dénaturation	95	50 s	
Hybridation	55	50 s	25
Elongation	72	50 s	
Elongation finale	72	10 min	

**Tableau 42 : Conditions expérimentales pour l'amplification par PCR des marqueurs STR13 et STR22.**

Les produits de PCR ont été analysés par électrophorèse capillaire sur un analyseur ABI3130 (Applied Biosystem).

##### 5. Analyse *in vitro* de l'épissage

Pour les conditions de culture des lignées cellulaires HeLa et Huh7, la production des plasmides minigènes, la transfection des cellules et l'étude des transcrits, se rapporter à la page 100 (Analyse *in vitro* de l'épissage).

##### **Production des minigènes pour l'exon 13 du gène *F8*.**

Pour la création d'un plasmide minigène de l'exon 13 du gène *F8*, un fragment de 1182 pb composé de 202 pb de l'intron 12, de l'exon 13 et de 770 pb de l'intron 13 a été amplifié à partir des ADN du patient 1 et d'un témoin sain. Les produits de PCR obtenus ont été insérés au niveau du site *NdeI* du vecteur pTB2. Le plasmide témoin (pTB2ex13<sub>WT</sub>) et le plasmide muté (pTB2ex13<sub>DEL</sub>) ont été contrôlés par PCR. Cinq mutations ponctuelles ont été produites sur le plasmide pTB2ex13<sub>WT</sub> par

mutagénèse dirigée au niveau de la queue poly-T de la séquence *AluY* inversée de l'intron 13 du gène *F8*. La séquence de ce nouveau plasmide (pTB2ex13<sub>WT\_5mut</sub>) a été vérifiée par séquençage.

### **Transfection cellulaire et analyse des transcrits.**

Les cellules HeLa et Huh7 ont étéensemencées dans des plaques 6 puits (100 000 cellules / puits). Le lendemain les cellules ont été transfectées avec 1µg de plasmide à l'aide du réactif JetPrime (Polypus transfection, Illkirch-Graffenstaden, France) en suivant les recommandations du fournisseur. Quarante-huit heures après transfection, les ARNm totaux ont été extraits puis transformés en ADNc par RT-PCR. Après amplification par PCR des produits de RT-PCR (couple d'amorces utilisé :  $\alpha$ 23G/ $\beta$ 2F), les amplicons ont été analysés par migration sur gel d'agarose et séquencés. Toutes les transfections ont été réalisées en triplicat.

#### **6. Inhibition de l'expression de l'hnRNP C**

### **Transfection des cellules HeLa**

Pour les expériences d'inhibition de l'expression de l'hnRNP C, le protocole de transfection des cellules HeLa a été modifié comme indiqué ci-dessous :

- Les cellules ont étéensemencées dans une plaque six puits le premier jour.
- Les cellules ont été transfectées le lendemain matin soit par un siRNA anti-hnRNP C (FlexiTube GeneSolution GS3183 for HNRPN C, Qiagen) soit par un siRNA contrôle (AllStar negative control siRNA, Qiagen) à la concentration finale de 50pM.
- Le jour 3, le milieu de culture a été changé et les cellules ont été transfectées avec 1µg soit de plasmide pTB2ex13<sub>WT</sub> soit de plasmide pTB2ex13<sub>DEL</sub>.
- Les ARNm ont été extraits et analysés comme précédemment le jour 5.

### **Western blot**

L'action du siRNA anti-hnRNP C sur la production d'hnRNP C par les cellules HeLa a été contrôlée par western blot sur les lysats de cellules. Ces derniers ont été préparés de la manière suivante : trois jours après transfection par le siRNA, les tapis cellulaires ont été lavés deux fois par du

DPBS 1X puis incubés 20 minutes dans la glace après ajout de 350 µl de tampon de lyse (HEPES pH 7,5 50mM, KCl 0,1M, MgCl<sub>2</sub> 2mM, Triton X-100 0,5%) contenant un inhibiteur de protéase (cOmplete Mini protease inhibitor). Après une étape de centrifugations pour éliminer les débris cellulaires, les lysats ont été conservés à -80°C.

Avant la réalisation du western blot, les échantillons ont été dénaturés par ajout d'un tampon de dénaturation et par chauffage à 95°C pendant 5 minutes. Ensuite, des volumes de lysat cellulaire correspondant à 20 µg de protéines totales ont été déposés dans les puits d'un gel (MiniPROTEAN TGX gel 12%, Biorad). Les protéines totales ont été séparées en fonction de leur taille par électrophorèse sur gel (130 V, 1h). Les protéines ont ensuite été transférées sur une membrane de difluorure de polyvinylidène. La membrane a été saturée pendant 1 heure (tampon TBS-Tween 1X / lait 5%) puis incubée une nuit à 4°C en présence de l'anticorps monoclonal de souris anti-hnRNP C dilué au 1/1000 (Anti-hnRNP C1+C2 [4F4], Abcam, Cambridge, UK). Le lendemain, après trois étapes de lavage dans du tampon TBS-tween, la membrane a été incubée 1 heure à température ambiante avec un anticorps polyclonal anti-anticorps de souris marqué par une peroxydase dilué au 1/10000 (GE Healthcare, Chicago, IL, Etats-Unis). Après trois nouveaux lavages en TBS-Tween, les bandes d'intérêt ont été révélées par chimiluminescence (ECL reagent, GE Healthcare) et impressionnées sur un film photographique (Fujifilm, Tokyo, Japon).

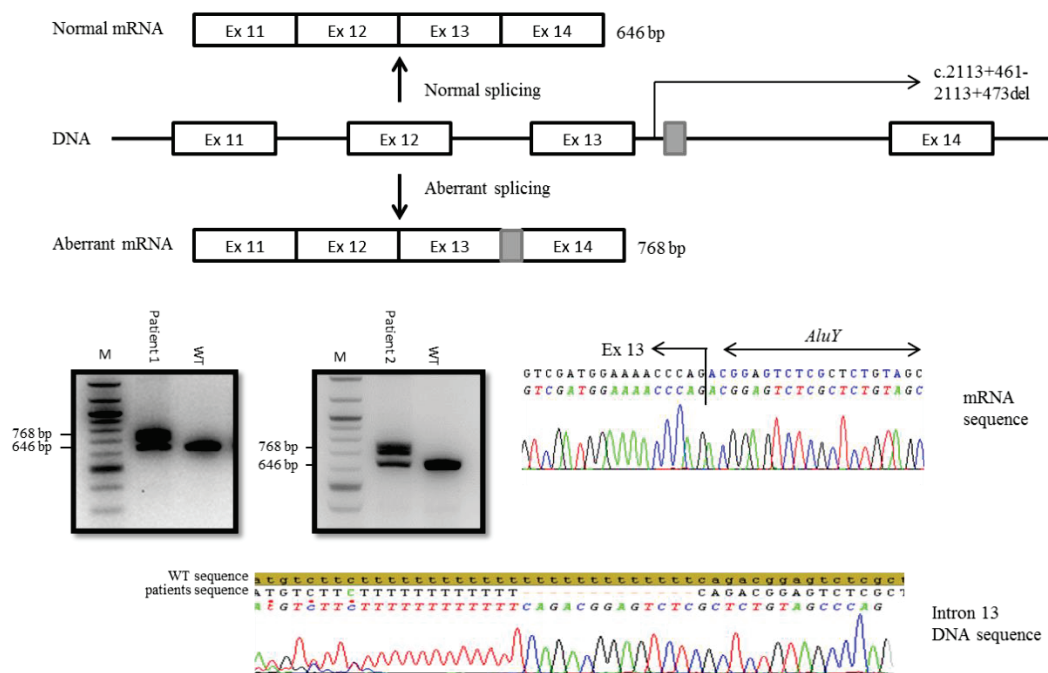
Un western blot de contrôle a été réalisé sur les mêmes échantillons dans les mêmes conditions mais en utilisant comme anticorps primaire un anticorps de lapin anti-GAPDH (rabbit mAb anti-GAPDH (D16H11) XP, Cell Signaling Technology, Danvers, MA, Etats-Unis).

## **5. Résultats**

**L'analyse des transcrits ectopiques du gène *F8* a montré l'insertion de la même séquence *AluY* entre les exons 13 et 14 chez les deux patients hémophiles A mineurs.**

La même anomalie a été mise en évidence chez les deux cas index lors de l'analyse des transcrits leucocytaires ectopiques du gène *F8*. Deux transcrits de taille différente ont été mis en évidence lors de l'analyse de la région s'étendant des exons 11 à 14 des transcrits du gène *F8* : le

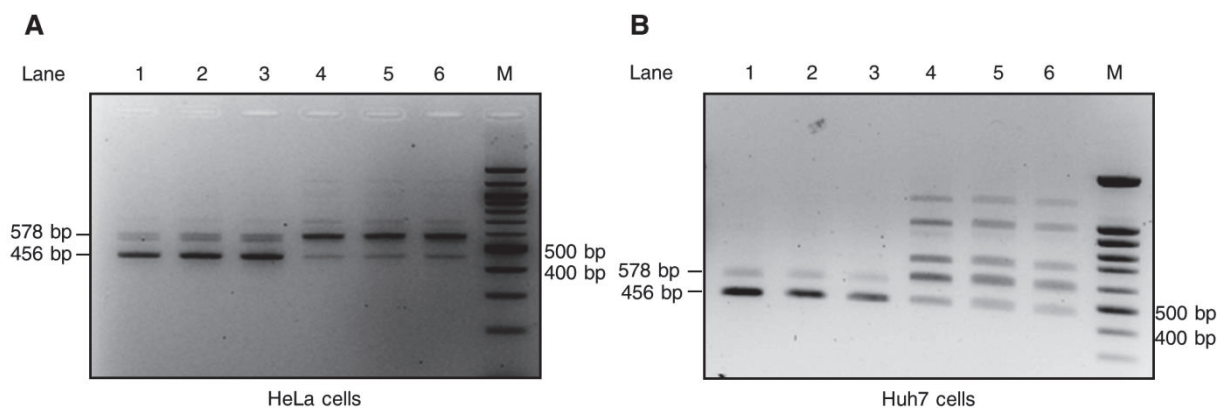
premier fait 646 pb et correspond au transcrit normal alors que le second est plus grand (768 pb) (Figure 42). Le séquençage du produit de 768 pb a montré l'insertion entre les exons 13 et 14 d'un fragment de 122 pb appartenant à l'intron 13 du gène *F8* (c.2113\_2114ins2113+477\_2113+598). Cette insertion est responsable d'un décalage du cadre de lecture (p.Gly705Aspfs\*37). Le séquençage de la région d'intérêt de l'intron 13 du gène *F8* a montré à la fois l'existence de séquences consensus d'épissage AG/GT de part et d'autre du pseudoexon et la présence d'une délétion de 13 pb (c.2113+461\_2113+473del) dans une répétition de 25 T juste en amont du pseudoexon. Cette délétion a été recherchée dans une population témoin non hémophile A et n'a pas été retrouvée (106 allèles étudiés). L'analyse de la séquence de ce pseudoexon a montré qu'il correspondait à une séquence *AluY* orientée en antisens par rapport au gène *F8* et que la délétion entraînait la perte d'une partie de la queue poly(T) de cette séquence *Alu*.



**Figure 43 : Identification de la délétion intronique c.2113+461\_2113+473del chez les deux cas index.** Deux transcrits différents ont été observés lors de l'analyse de la région s'étendant des exons 11 à 14 (transcrit normal : 646 pb et transcrit aberrant : 768 pb). Le séquençage du fragment de 768 pb a montré la présence d'une insertion d'un fragment de 122 pb appartenant à l'intron 13 du gène *F8* (c.2113\_2114ins2113+477\_2113+598). Le séquençage de l'intron 13 a mis en évidence l'existence d'une petite délétion intronique (c.2113+461\_2113+473del) dans la queue poly-(T) d'une séquence *AluY* orientée en antisens par rapport au *F8*. M, marqueur de taille ; WT : témoin normal ; Ex, exon.

**Le rôle de la délétion c.2113+461\_2113+473del dans l'exonisation de la séquence *AluY* est confirmé par analyse fonctionnelle *in vitro*.**

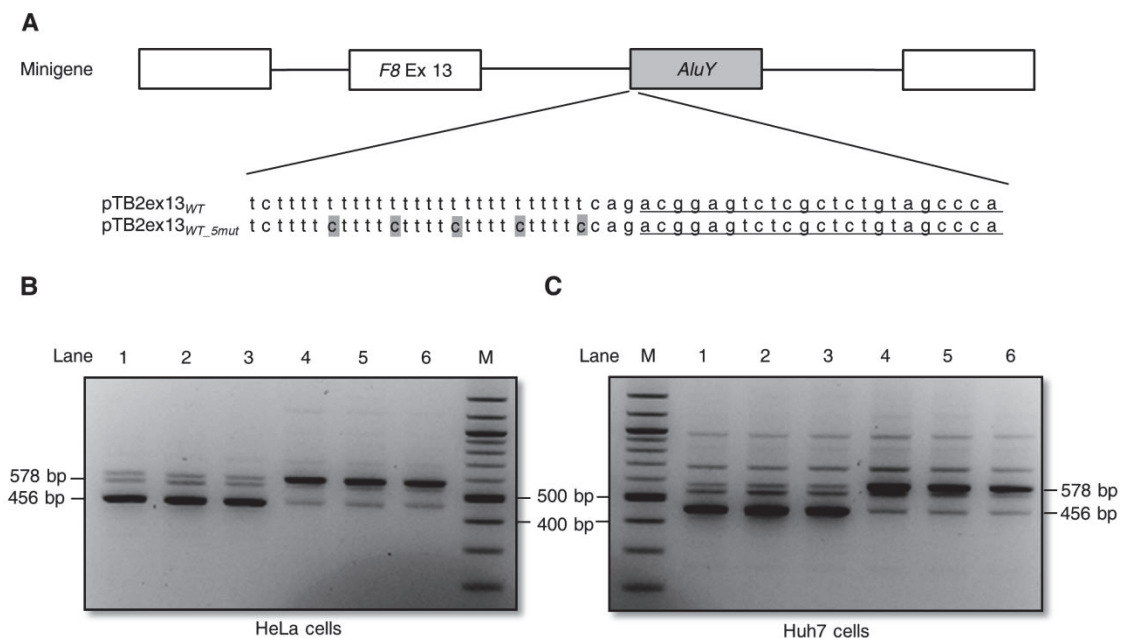
Afin de démontrer le lien de cause à effet entre la délétion c.2113+461\_2113+473del et l'exonisation de la séquence *AluY*, une analyse fonctionnelle *in vitro* par technique minigène a été réalisée. Deux minigènes contenant l'exon 13 du gène *F8* ont été produits, l'un correspondant à la séquence naturelle (pTB2ex13<sub>WT</sub>) et l'autre à la séquence mutée (pTB2ex13<sub>DEL</sub>), et transfectés dans des cellules HeLa. Les transcrits des minigènes ont été étudiés 48h après transfection (Figure 44 A). Une bande majoritaire de 456 pb correspondant à un épissage correct de l'exon 13 du gène *F8* est observée pour les cellules transfectées par le minigène pTB2ex13<sub>WT</sub>. On note aussi la présence de trace de transcrits de 578 pb correspondant à l'insertion de la séquence *AluY*. A l'inverse, dans le cas des cellules transfectées par le minigène pTB2ex13<sub>DEL</sub>, le transcrit de 578 pb est majoritaire et on observe uniquement des traces de transcrits normaux. Afin de confirmer cette observation, la même expérience a été réalisée dans une autre lignée cellulaire (Huh7) et les mêmes résultats ont été obtenus (Figure 44 B).



**Figure 44 : Résultats de l'analyse fonctionnelle *in vitro* de l'impact de la délétion c.2113+461\_2113+473del sur l'exonisation de la séquence *AluY*.** Les cellules HeLa (A) et Huh7 (B) ont été transitoirement transfectées soit avec le plasmide minigène non muté (pTB2ex13<sub>WT</sub>; puits 1 à 3) soit avec le plasmide minigène muté (pTB2ex13<sub>DEL</sub>; puits 4 à 6). Quarantes-huit heures après transfection, ARNm ont été analysés par RT-PCR (épissage normal de l'exon 13 : 456 pb ; insertion de la séquence *AluY* : 578 pb). M : marqueur de taille.

**L'insertion de mutation dans la queue poly-(T) de la cette séquence *Alu* et l'inhibition de la synthèse d'hnRNP C dans les cellules reproduit l'effet de la délétion c.2113+461\_2113+473del.**

Afin d'étudier le rôle de l'hnRNP C sur la répression de la séquence *AluY* de l'intron 13 du gène *F8*, deux expériences *in vitro* ont été réalisées. Dans la première, la queue poly-(T) de la séquence *Alu*, zone de fixation de l'hnRNP C, a été discontinuée par introduction de 5 mutations (T>C) sur le plasmide pTB2ex13<sub>WT\_5mut</sub>. L'analyse des transcrits des cellules HeLa transfectées avec ce plasmide a mis en évidence que les mutations présentes au niveau de la queue poly-(T) favorisaient l'exonisation de la séquence *Alu* (Figure 45).

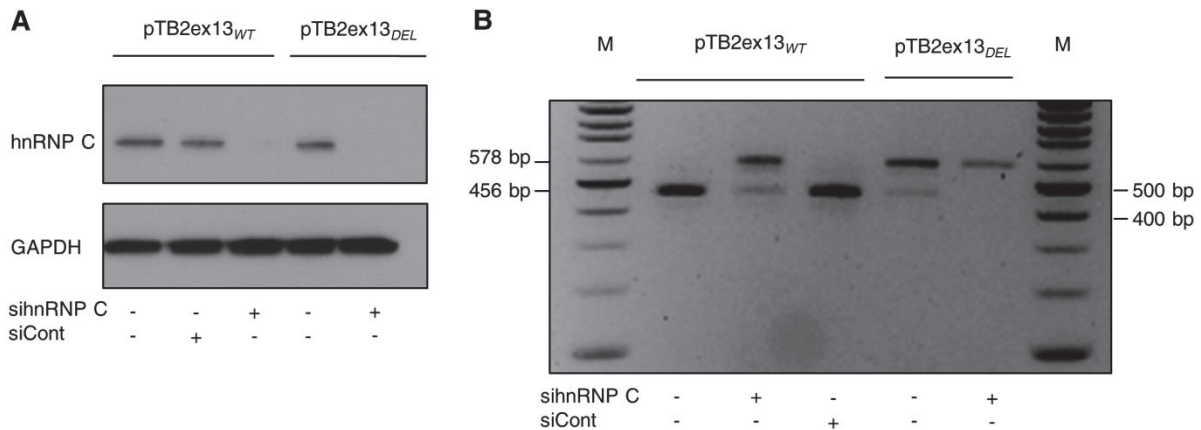


**Figure 45 : L'insertion de mutation dans la queue poly-(T) de la séquence *AluY* favorise son exonisation.** (A) Représentation schématique des constructions minigènes avec la séquence *AluY* (rectangle gris), l'exon 13 (rectangle blanc) et les régions introniques (traits noirs). Les différences entre la séquence des plasmides pTB2ex13<sub>WT</sub> et pTB2ex13<sub>WT\_5mut</sub> sont indiquées en gris. Le début de la séquence *AluY* exonisée est souligné. (B) Résultats de l'analyse des transcrits des cellules HeLa 48 heures après transfection par les plasmides pTB2ex13<sub>WT</sub> (puits 1 à 3) ou pTB2ex13<sub>WT\_5mut</sub> (puits 4 à 6). (C) Résultats de l'analyse des transcrits des cellules Huh7 48 heures après transfection par les plasmides pTB2ex13<sub>WT</sub> (puits 1 à 3) ou pTB2ex13<sub>WT\_5mut</sub> (puits 4 à 6). M, marqueur de taille.

Le rôle de la liaison de l'hnRNP C dans la répression de la séquence *AluY* a ensuite été confirmé par des expériences dans lesquelles la production de cette protéine dans les cellules HeLa était inhibée à l'aide d'un siRNA spécifique. Dans un premier temps, l'efficacité du siRNA anti-hnRNP C a été vérifiée par western blot (Figure 46 A). L'hnRNP C, présente de manière constitutive dans les lysats de cellules, devient indétectable quand les cellules sont transfectées par le siRNA anti-



hnRNP C. L'inhibition de la production de hnRNP C dans les cellules HeLa reproduit l'effet de la mutation c.2113+461\_2113+473del en favorisant l'exonisation de la séquence *AluY* (Figure 46 B). Ces résultats suggèrent fortement que la répression de l'exonisation de cette séquence *AluY* par l'hnRNP C est diminuée par la délétion c.2113+461\_2113+473del.



**Figure 46 : L'inhibition de la production d'hnRNP C favorise l'exonisation de la séquence *AluY* dans les cellules HeLa.** (A) Western blots réalisés sur les lysats cellulaires pour vérifier l'efficacité du siRNA anti-hnRNP C. (B) Résultats de l'analyse des transcrits des cellules HeLa traitées par siRNA et transfectées par les plasmides pTB2ex13<sub>WT</sub> et pTB2ex13<sub>DEL</sub>. M, marqueur de taille.

### La délétion c.2113+461\_2113+473del est une mutation récurrente dans la population française d'hémophiles A mineurs.

La mutation c.2113+461\_2113+473 a été recherchée dans toutes les familles avec antécédent d'HA mineure pour lesquelles aucune anomalie génétique n'était détectée à ce jour. Parmi les 1970 familles d'hémophiles A connues dans le laboratoire de biologie moléculaire de Lyon, 992 (50.4%) présentent un phénotype mineur d'HA. Aucune mutation n'avait été trouvée dans 114 de ces familles (114/992 = 11,5%). Une délétion (de 11 à 13 pb) dans la queue poly-(T) de la séquence *AluY* de l'intron 13 du gène *F8* a été observée chez 61 de ces familles (61/114 = 54%). La délétion c.2113+461\_2113+473 (13 pb) a été trouvée chez 57 familles, la délétion c.2113+462\_2113+473 (12 pb) chez 1 famille et la délétion c.2113+463\_2113+473 (11 pb) chez 3 familles. Ainsi, les petites délétions présentes dans cette région de l'intron 13 sont responsables de 6,1% des hémophilies A

mineurs dans notre cohorte. Les patients hémophiles porteurs de ces anomalies ont des taux de FVIII:C proches de 30% (moyenne +/- écart type : 30 /+ 6%).

Devant la haute prévalence de cette anomalie génétique chez les hémophiles A mineurs, l'hypothèse de l'existence d'un effet fondateur a été émise. Afin de vérifier cette hypothèse, une analyse des polymorphismes intragéniques du gène *F8* a été réalisée chez 29 patients porteurs de la délétion c.2113+461\_2113+473 et chez tous les patients porteurs des délétions de 11 et 12 pb (Tableau 43). Un même haplotype caractérisé par les allèles suivants : 248 (STR1), 201 (STR9), 149 (STR13), 178 (STR22), 134 (STR25A) et 255 (STR25B) a été identifié chez la totalité des patients porteurs de la délétion c.2113+461\_2113+473del. Ces observations sont en faveur de l'existence d'un effet fondateur dans la population française pour cette délétion. De façon intéressante, 4 haplotypes différents ont été trouvés chez les 4 patients porteurs de délétions de 11 et 12 pb.

Délétion	STR 1	STR 9	STR 13	STR 22	STR 25A	STR 25B	Fréquence, n (%)
c.2113+461_2113+473del	248	201	149	178	134	255	29 <sup>a</sup> (100)
c.2113+461_2113+471del	246	201	147	180	134	257	1 (33)
	250	201	153	172	134	253	1 (33)
	246	201	145	180	134	257	1 (33)
c.2113+461_2113+472del	246	201	145	180	134	257	1 (100)

**Tableau 43 : Analyse des polymorphismes intragénique du gène *F8* chez les patients porteurs de délétion dans la queue poly-(T) de la séquence AluY de l'intron 13.** <sup>a</sup>L'analyse des polymorphismes intragénique du gène *F8* a été réalisée uniquement chez 29 patients porteurs de la délétion de 13 pb.

## 6. Discussion

Dans ce travail, des analyses de transcrits ectopiques leucocytaires du gène *F8* ont été réalisées chez deux hémophiles A mineurs pour lesquels aucune anomalie génétique n'était connue. La même insertion d'une séquence intronique de 122 pb (c.2113\_2114ins2113+477\_211+598) entre les exons 13 et 14 a été observée dans les ARNm de ces deux patients. L'analyse des transcrits réalisée chez ces patients a aussi permis de mettre en évidence la co-existence du transcrit normal avec le transcrit aberrant, ce qui est en accord avec le phénotype clinique observé chez ces patients. Le séquençage de l'intron 13 du gène *F8* a mis en évidence la même petite délétion de 13 pb chez ces deux patients (c.2113+461\_2113+473del) localisée quelques paires de bases en amont du pseudoexon. La pathogénicité de cette délétion est confortée à la fois pour les études *in vitro* réalisées dans les cellules HeLa et Huh7 ainsi que par l'absence de détection de cette anomalie chez 55 sujets contrôles (106 allèles différents). Cette étude est la première description d'une délétion intronique profonde

entraînant l'activation d'un exon cryptique dans l'hémophilie A. En effet, les quelques activations d'exons cryptiques décrites à ce jour dans l'hémophilie A étaient toutes la conséquence de mutations ponctuelles introniques profondes. Ces mutations entraînaient la création de sites d'épissage accepteurs ou donneurs *de novo* (c.143+1567A>G, c.5587-93C>T et c.1537+325A>G) ou bien augmentaient la force de sites cryptiques d'épissage pré-existant (c.5999-277G>A, c.5998+941G>A et c.2113+1152delA) (58,123–125).

L'étude de la séquence des 122 pb exonisée dans les transcrits matures de ces deux patients a révélé que ce pseudoexon correspondait à une séquence *AluY* orientée en antisens. Les séquences *Alu* peuvent s'insérer au sein de la séquence codante d'un gène par deux mécanismes principaux. Le premier est la rétrotransposition *de novo* d'une séquence *Alu* au sein d'un exon. Ce phénomène a été décrit à deux reprises chez des hémophiles A sévères (218,254). Le second mécanisme consiste en l'exonisation de la séquence *Alu*, comme cela est le cas dans cette étude. Plus de 90% des séquences *Alu* exonisées responsables de pathologies génétiques sont orientées en antisens (258). Ceci est probablement lié au fait que les séquences *Alu* orientées en antisens possèdent des structures très proches d'un site accepteur d'épissage (253). Dans notre cas, la séquence *AluY* orientée en antisens et localisée dans l'intron 13 du gène *F8* est épissée au niveau du site accepteur de la position *Alu* 279 (site utilisé par 24% des séquences *Alu* exonisées physiologiquement chez l'Homme) et du site donneur de la position *Alu* 157 (site utilisé par 65% des séquences *Alu* exonisées physiologiquement chez l'Homme) (264,265,271).

La délétion c.2113+461\_2113+473del entraîne la perte d'une partie de la queue poly-(T) de la séquence *AluY* de l'intron 13 du gène *F8*. Récemment, il a été montré que la queue poly-(T) des séquences *Alu* antisens était le site de fixation de l'hnRNP C, une protéine ayant pour fonction la protection du génome contre l'exonisation de ces séquences *Alu* (266,267). Nous avons émis l'hypothèse que la délétion c.2113+461\_2113+473del pourrait diminuer la fixation de l'hnRNP C et donc entraîner l'exonisation de la séquence *AluY*. Nos analyses *in vitro* ont montré que l'insertion de mutations dans la queue poly-(T) de la séquence *AluY* (élimination des sites de fixation de l'hnRNP C) et l'inhibition de la synthèse d'hnRNP C à l'aide d'un siRNA reproduisaient l'effet de la délétion

c.2113+461\_2113+473del. Ces résultats confortent grandement notre hypothèse initiale. Un mécanisme moléculaire similaire a déjà été décrit dans la littérature chez un patient ayant une hyperphénylalaninémie chez lequel la délétion totale de la queue poly-(T) d'une séquence *AluS<sub>q</sub>* antisens dans l'intron 2 du gène *PTS* entraînait une exonisation totale de cette séquence *AluS<sub>q</sub>* dans les transcrits matures (266,272). Dans les deux cas index présentés ici, il est intéressant de remarquer que la délétion partielle de la queue poly-(T) est suffisante pour avoir un effet délétère significatif sur la production de FVIII.

De façon surprenante, des délétions de 11 à 13 pb dans la queue poly-(T) de la séquence *AluY* de l'intron 13 du gène *F8* ont été retrouvées dans plus de la moitié des familles d'hémophiles A mineurs pour lesquelles aucune anomalie génétique n'était connue à ce jour. Avec une prévalence de 6,1 %, cette mutation est la plus fréquemment retrouvée chez les hémophiles A mineurs dans notre population. Les mutations retrouvées avec des fréquences élevées dans une population sont la conséquence soit d'un effet fondateur soit d'un mécanisme moléculaire récurrent (comme cela est le cas pour les Inv22 et Inv1). Des effets fondateurs ont déjà été décrits pour plusieurs mutations trouvées dans des populations isolées d'hémophiles A (270,273,274). L'analyse des polymorphismes intragéniques du gène *F8* a montré que tous les patients porteurs de la mutation c.2113+461\_2113+473del avaient le même haplotype. Cette observation est en faveur d'un effet fondateur pour cette délétion. Cependant le fait que les patients portant les mutations c.2113+462\_2113+473del et c.2113+463\_2113+473del ont des haplotypes différents pourrait indiquer l'existence d'un mécanisme moléculaire récurrent. Ainsi il serait intéressant de rechercher ces anomalies dans d'autres populations d'hémophiles A mineurs.

En conclusion, dans cette étude nous avons décrit un nouveau mécanisme moléculaire responsable d'HA mineure lié à une délétion de 13 T (c.2113+461\_2113+473del) dans la queue poly-(T) d'une séquence *AluY* antisens. Du fait de la haute prévalence de ce type d'anomalie dans la population française d'hémophiles A, cette délétion doit être spécifiquement recherchée chez tous les patients hémophiles A mineurs chez lesquels aucune anomalie génétique n'est mise en évidence lors de l'analyse moléculaire de routine du gène *F8*.

## 7. Annexes

### 1. Séquence des amorces utilisées pour l'étude des transcrits ectopiques

Couple	Amorces sens		Amorces antisens		Taille du fragment, pb
	Exon	Séquence	Exon	Séquence	
A	1	CTTTTGGCGATTCTGCTTTAG	9	GCTGAGGGCCATTGTTCAAA	1273
A1	1	CTTTTGGCGATTCTGCTTTAG	5-6	GAGTGCCAACTTTTCCCTTCAT	650
A2	5	AAGGAAAAGACACAGACCTT	7-8	TCATGTTGGTGGGAAGAGATA	398
B	7	TACTGCTCAAACACTCTTG	14	TCATATTTGGCTTCTGGAG	1561
B1	7	TACTGCTCAAACACTCTTG	11	AGTTGGCCCATCTTCTACAGTCACTGTC	686
B2 <sup>a</sup>	13	TGCCTGACCCGCTATTACT	14	AGAAGCTTCTTGGTTCAATG	645
C	14	GATACCATTTTGTCCCTGAA	22	ACATGATGATAAACTGAGAGA	1528
C1	14	GATACCATTTTGTCCCTGAA	17	CATGGAAGCGATAAT	973
C2	17	AAAAATATGGAAAGAACTGC	22	ACATGATGATAAACTGAGAGA	636
D	19	CCTTATTGGCGAGCATCTACA	3'UTR	TTGCCATAGTTATATTGGAAG	1254
D1	22	TTATTCACGGCATCAAGACCC	23	AAATCACAGCCCATCAACTCC	274
D2	23	CTCCAATTATTGCTCGATACAT	25	CTGGTAAGCAGAGATTTTAC	337
D3 <sup>a</sup>	24	GCCATTGGGAATGGAGAGTA	3'UTR	AGTTAATTCAGGAGGCTTCA	606

Les amorces antisens des couples A, B, C et D ont été utilisés pour la réaction de RT. Les couples d'amorces A, B, C et D ont été utilisés pour la première réaction d'amplification. Les couples d'amorces A1, A2, B1, B2, C1, C2, D1, D2, D3 ont été utilisés pour la seconde réaction d'amplification. <sup>a</sup> Amorces publiées par El-Maarri *et al* (209).

### 2. Séquence des amorces utilisées pour le séquençage de l'intron 13 du gène F8 et pour la construction des minigènes.

Couple	Amorces sens		Amorces antisens		Taille du fragment, pb
	Position	Séquence	Position	Séquence	
E	Intron 12 (-202)	TGGTTTATGACTGTCTCCTCACA	Intron 13 (+770)	GCTGAGGGCCATTGTTCAAA	1182
F	Exon 13	AGCATTGGAGCACAGACTGA	Intron 13 (+675)	AGACCATCCTGGCTAACACG	810
G	Intron 12 (-202)	<i>CTAAACAGCCACATATGTGGTTTA</i> TGACTGTCTCCTCACA	Intron 13 (+770)	<i>CCCCCTCGACCATATGCAAAAGAA</i> GACATTCAGGCTGGGC	1182

Le couple d'amorce E a été utilisé pour l'amplification de la région de l'intron 13 d'intérêt. Le couple F a été utilisé pour la réaction de séquençage. Le couple G a été utilisé pour la création des minigènes. Les séquences en italique correspondent aux queues spécifiques nécessaires pour la réaction in-fusion.

### 3. Séquence des amorces utilisées pour l'étude des microsatellites

Primer pair	Sequence of forward primer	Sequence of reverse primer
STR1	TTGCAAATGTACAATAGGGCAGT	TTCTTGGTCTGCCTTCTGA
STR9	TGGACTTCCAACCCCATAGTC	CACCATGCCTGGCTAATTCA
STR13	TGCATCACTGTACATATGTATCTT	CCAAATTACATATGAATAAGCC
STR22	TTCTAAGAATGTAGTGTGTGTG	TAATGCCACATTATAGA
STR25A	AGATCGCGCCATCACATTC	AGGGGTAGGCAGGCTTGTTT
STR25B	CAGAACCAATTCCAGAAATCCA	GCTCTCTGGGGTCTCTAGGC
F8_del_IVS13	CTGGAGTGAAATATCTCATTGTGC	AAAATTAGCTGGGTGTGGTGG

Les amorces antisens sont marquées par un fluorochrome (FAM).

**Partie2 : Etude des effets fonctionnels de la mutation c.1611C>A (p.Cys537X) du gène *THBD*.**

**Why patients with THBD c.1611C>A (p.Cys537X) nonsense mutation have high level of soluble thrombomodulin ?**

**Yohann Jourdy**\*†, Nathalie Enjolras †, Sandra Le Quellec \*†¶, Jean Claude Bordet \*†, Claude Négrier \*†¶, Christine Vinciguerra \*†, Yesim Dargaud \*†¶

PlosOne 2017, 12 :e0188213.

\* Hospices Civils de Lyon, Centre de Biologie et Pathologies Est, Service d'hématologie Biologique, Bron, France

† EAM 4609 Hémostase et cancer, Université Claude Bernard Lyon 1, France.

¶ Hospices Civils de Lyon, Hôpital Cardiologique Louis Pradel, Unité d'Hémostase Clinique, Bron, France.

**Keywords:** THBD c.1611C>A, thrombomodulin, proteases, reducing agents, shear stress

## 1. Résumé

**Contexte :** Notre équipe a récemment décrit une nouvelle pathologie hémorragique caractérisée par la présence d'une concentration très élevée de thrombomoduline soluble dans le plasma chez une patiente. Cette pathologie se transmet sous le mode autosomique dominant. La mutation du gène *THBD* c.1611C>A (p.Cys537X) a été mise en évidence à l'état hétérozygote chez le propositus. Cette mutation entraîne la synthèse d'une protéine plus courte dans laquelle les 3 derniers acides aminés du domaine transmembranaire ainsi que le domaine intra-cytoplasmique sont manquants.

**Objectif :** Déterminer par quels mécanismes la thrombomoduline mutée (p.Cys537X) est libérée dans la circulation sanguine.

**Matériels et méthodes :** un plasmide contenant l'ADNc du gène *THBD* (TM-WT) a été muté par mutagenèse dirigée pour créer la mutation c.1611C>A (TM<sub>1-536</sub>) ainsi que différents mutants de la thrombomoduline (TM<sub>1-515</sub>, TM<sub>1-525</sub>, TM<sub>1-529</sub> et TM<sub>1-537</sub>) possédant des domaines transmembranaires de taille variable. Des cellules COS-1 ont été transitoirement transfectées avec ces plasmides. L'effet des forces de cisaillement, de certaines protéases et de certains agents réducteurs sur la libération de thrombomoduline dans les surnageants cellulaires a été évalué.

**Résultats :** les analyses par western-blot et par immunofluorescence ont montré que la TM<sub>1-536</sub> était synthétisée par les cellules COS-1 et qu'une certaine fraction était présente à la surface des cellules. Un taux de thrombomoduline soluble significativement plus élevé a été observé dans les surnageants de cellules exprimant la TM<sub>1-536</sub> en comparaison avec les cellules exprimant la TM-WT (respectivement 56,3 vs 8,8 ng/ml,  $p=0,001$ ). La libération de la TM<sub>1-536</sub> dans les surnageants cellulaires était diminuée de 75% lorsque les cellules étaient cultivées en présence d'inhibiteurs de métalloprotéases. Aucune différence sur la libération de TM<sub>1-536</sub> et TM-WT n'a été observée après exposition des cellules à différentes concentrations de cathépsine, élastase et de plusieurs agents réducteurs ainsi qu'à d'importantes forces de cisaillement ( $5000\text{s}^{-1}$ ). Un taux significativement plus élevé de thrombomoduline a été observé dans les surnageants de cellules exprimant les TM<sub>1-515</sub>, TM<sub>1-525</sub>, TM<sub>1-529</sub> par rapport à la TM-WT ( $p<0,05$ )



**Conclusion** : le mécanisme moléculaire responsable de la libération de la TM<sub>1-536</sub> est complexe et non complètement élucidé : une plus grande sensibilité au clivage par les métalloprotéases ainsi qu'une anomalie de biosynthèse liée à la diminution de la taille du domaine transmembranaire pourraient expliquer les très haut taux de thrombomoduline soluble observés chez les patients porteurs de la mutation c.1611C>A.

## 1. Rappels bibliographiques sur la thrombomoduline

La thrombomoduline (TM, CD141) est une protéine transmembranaire qui a été décrite la première fois en 1981 en tant que cofacteur de la thrombine pour l'activation de la protéine C (PC) (275). Depuis sa découverte, il a été démontré que cette protéine intervenait dans de nombreux processus physiologiques et pathologiques.

### 1. Expression de la thrombomoduline

La thrombomoduline est présente à la surface des cellules endothéliales de tous les vaisseaux sanguins (artères, veines, capillaires) et lymphatiques. Elle est aussi exprimée au niveau du syncytiotrophoblaste (276). De nombreuses autres cellules expriment la TM à leur surface, parmi lesquelles : les kératinocytes, les astrocytes, les cellules mésothéliales, les ostéoblastes, les monocytes, les chondrocytes et les cellules épithéliales alvéolaires (277).

### 2. Structure de la thrombomoduline

La TM est codée par le gène *THBD* localisé sur le chromosome 20 (20p11.21). Ce gène s'étend sur environ 4kb et est composé d'un unique exon (absence d'intron). La TM est synthétisée sous la forme d'une chaîne peptidique de 575 acides aminés (aa). Après l'élimination du peptide signal (18 aa), la protéine mature est constituée d'une chaîne polypeptidique de 557 aa. Cette protéine est divisée en 5 domaines (TMD) distincts (Figure 47).

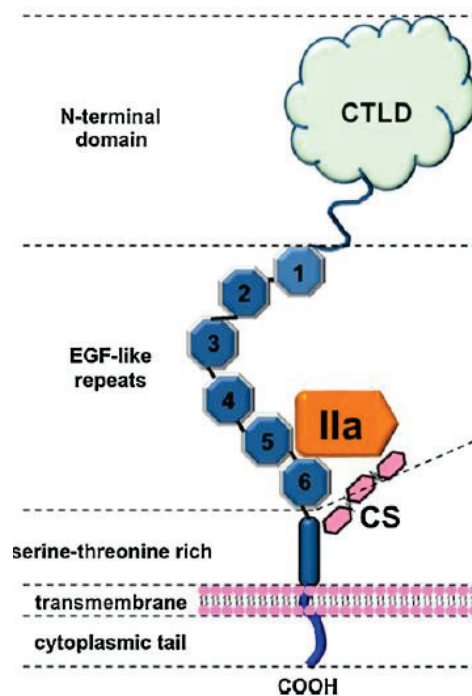
La région N-terminale de la TM (résidus 1-222, TMD1) contient une région (résidus 1-155) présentant une très forte homologie avec les lectines de type C. Néanmoins le site de liaison au  $Ca^{2+}$  et le domaine de reconnaissance des sucres, caractéristique des lectines, sont absents. Ce domaine est globulaire et est situé de manière idéale (au contact du plasma) pour interagir avec de nombreuses molécules en circulation. Ce domaine n'a aucune fonction anticoagulante mais joue un rôle important dans l'inflammation, l'adhésion cellulaire et la prolifération cellulaire.

Le domaine suivant (TMD2) est composé de 6 répétitions de type EGF (epidermal growth factor). Les fonctions des deux modules EGF en position C-terminale (EGF 1 et 2) sont inconnues. Les 4

autres répétitions (EGF 3, 4, 5 et 6) sont nécessaires à l'activation de la PC et de l'inhibiteur de la fibrinolyse activable par la thrombine (TAFI).

Le troisième domaine de la TM (TMD3) est une région riche en sérines et thréonines contenant 4 sites potentiels d'O-glycosylation. Une chaîne de sulfate de chondroïtine (chaîne composée d'environ une vingtaine de disaccharides et terminée par un trisaccharide) est fixée dans cette région (278).

Le quatrième domaine (TMD4) est le domaine transmembranaire et le domaine C-terminal de la TM (TMD5) est intracellulaire. Ce domaine comprend une seule cystéine qui est impliquée dans la polymérisation de la molécule (279). Ce domaine n'est pas très conservé entre les espèces.



**Figure 47 : structure de la thrombomoduline (277).** CTLD : domaine de lectine de type C ; EGF : domaine de type epidermal growth facteur ; CS : chaîne de sulfate de chondroïtine.

### 3. La thrombomoduline soluble

Une forme soluble de la TM (sTM) est présente dans le plasma et l'urine (280). La sTM est produite à partir de la TM présente à la surface des cellules endothéliales par clivage chimique ou enzymatique (281–283) souvent à la suite de dommages vasculaires liés à un contexte inflammatoire (infection, sepsis). Le taux de sTM dans le plasma est très faible (proche de 10 ng/ml) chez un

individu sain (280). Le taux de sTM circulant augmente significativement dans de nombreuses pathologies entraînant des lésions vasculaires (coagulation intravasculaire disséminée, purpura thrombotique thrombocytopénique, microangiopathie diabétique) (284) c'est pourquoi le taux de sTM est parfois utilisé comme marqueur de lésion vasculaire (285).

#### 4. Fonctions biologiques de la thrombomoduline

##### a) Propriétés anticoagulantes

Les fonctions anticoagulantes de la TM sont liées à la formation du complexe TM/thrombine. La TM exprimée à la surface des cellules endothéliales peut former un complexe de haute affinité avec la thrombine ( $K_d = 0,5 \text{ nM}$ ). Du fait de la très grande abondance de la TM dans la microcirculation, la majorité de la thrombine formée dans la circulation sanguine en dehors d'une brèche vasculaire est liée à la TM. Au sein du complexe TM/thrombine, la thrombine perd ses capacités :

- d'activation des plaquettes : inhibition de l'interaction entre la thrombine et le récepteur plaquettaire de la thrombine (PAR-1)
- de transformation du fibrinogène en fibrine
- d'activation des co-facteurs V et VIII.

A l'inverse sa capacité à activer la PC est exacerbée ( $\times 1000$ ). Du fait de la grande surface endothéliale présente dans les capillaires, l'activation de la PC dans les capillaires est très efficace. A l'inverse, dans les vaisseaux de plus gros calibre, le rapport surface endothéliale/volume sanguin est plus faible et la liaison de la PC à son récepteur endothélial (EPCR) permet d'augmenter l'activation de la PC par le complexe TM/thrombine par un facteur 6 à 20 (286,287). L'activation de la PC est aussi amplifiée par un facteur 25 par le facteur 4 plaquettaire (PF4) sécrété par les plaquettes activées. Le PF4 se lie au niveau du domaine Gla de la PC et favorise la liaison et l'activation de la PC par le complexe TM/thrombine (288). La PC activée (PCa), en association avec son co-facteur la protéine S, inhibe la boucle d'amplification de la thrombine en dégradant par clivage protéolytique les co-facteurs Va et VIIIa.

L'activité anticoagulante de la TM est portée par les domaines TMD2 et TMD3. La région minimale pour l'activation de la PC comprend les domaines EGF 4, 5 et 6. Les domaines EGF 5 et 6 interviennent dans la formation du complexe TM/thrombine (289). L'activation de la PC requiert les domaines EGF 4 et TMD3. La présence des groupes chondroïtine sulfate au niveau du TMD3 favorise l'activation de la PC par le complexe TM/thrombine (290). Ceci pourrait être lié à l'interaction entre les groupes chondroïtine sulfate et l'exosite I de thrombine (région affine pour les molécules chargées négativement).

#### b) Propriétés anti-fibrinolytiques

L'activation du TAFI par la thrombine est multipliée par 3 en présence de TM (291). Il a été montré que la fraction soluble de la TM participe aussi à l'activité anti-fibrinolytique de la TM (292). Le TAFI activé (TAFIa) clive les résidus lysines et arginines C-terminaux de la fibrine. L'élimination de ces résidus diminue la capacité de la fibrine à catalyser l'activation du plasminogène conférant ainsi au caillot de fibrine une certaine résistance à la fibrinolyse (293). Les régions EGF3, 4, 5 et 6 de la TM sont nécessaires à l'activation du TAFI par le complexe TM/thrombine. A l'inverse de son action sur la génération de PCa, le PF4 inhibe la génération de TAFIa par le complexe TM/thrombine (294). Une autre fonction anti-fibrinolytique du complexe TM/thrombine est la dégradation de l'activateur du plasminogène de type urokinase (uPA) (295–298).

#### c) Propriétés anti-inflammatoires

Le système de la PC est un acteur important de la réponse inflammatoire de l'hôte contre les infections (299,300). La PC activée par le complexe TM/thrombine peut moduler la réponse des cellules endothéliales et des cellules immunitaires à différents stimulus inflammatoires. Ainsi, la PCa peut réguler la production de cytokine par les monocytes, inhiber les voies de signalisation pro-inflammatoires dans les cellules endothéliales et entraîner l'activation de gènes cytoprotecteurs dans les endothéliums (301–303). Deux mécanismes moléculaires ont été décrits pour expliquer les propriétés anti-inflammatoires de la PCa. Le premier consiste en une activation des récepteurs PAR-1 présents à la surface des cellules endothéliales par le complexe PCa/EPCR. La stimulation du récepteur PAR-1 active des voies de signalisation aboutissant à l'activation de gènes anti-

inflammatoires et anti-apoptiques (304). Le second mécanisme est indépendant des récepteurs PAR-1 et implique la translocation au niveau du noyau des cellules du complexe PCa/EPCR entraînant l'activation de différents gènes aux propriétés anti-inflammatoires et cytoprotectives (305).

En plus de ces propriétés anti-inflammatoires indirectes liées à l'activation de la PC par le complexe TM/thrombine, la TM possède aussi une activité anti-inflammatoire directe via son domaine TMD1 (domaine N-terminal de type lectine) (306). Ce domaine a un rôle de protection de l'endothélium vasculaire contre les stimuli pro-inflammatoires et diminue la liaison des cellules immunitaires au niveau de l'endothélium notamment en inhibant l'expression de molécules d'adhésion telles que ICAM-1.

L'expression de la TM à la surface des cellules endothéliales est diminuée dans de nombreuses pathologies associées à une forte réaction inflammatoire telles que les sepsis et l'athérosclérose (307). Plusieurs mécanismes sont impliqués dans cette diminution de l'expression de la TM en contexte inflammatoire : l'inhibition de la transcription du gène de la TM induit par certaines cytokines, l'oxydation du domaine extracellulaire de la TM ou la libération de sTM à partir de la TM membranaire par clivage protéolytique par certaines enzymes (élastase et cathepsine) (281,308–310).

#### d) Propriétés anti-prolifératives

Des études ont montré que la TM avait une activité anti-proliférative sur les cellules musculaires lisses et les cellules endothéliales (311,312). Il a aussi été montré que le niveau d'expression de la TM à la surface de cellules tumorales était inversement proportionnel à leur prolifération (313).

## 2. Introduction

Récemment, plusieurs équipes indépendantes ont décrit une nouvelle pathologie hémorragique chez des patients porteurs à l'état hétérozygote de la mutation du gène *THBD* c.1611C>A (p.Cys537X). (130–133). Cette pathologie se transmet sous le mode autosomique dominant. Cette mutation entraîne la synthèse d'une protéine tronquée des 3 derniers aa du domaine transmembranaire et de la totalité domaine intracytoplasmique de la TM.

L'objectif de ce travail a été de déterminer par quels mécanismes la TM est libérée dans la circulation sanguine chez les patients porteurs de la mutation *THBD* c.1611C>A.

## 3. Matériels et méthodes

### 1. Vecteurs plasmidiques

Plusieurs constructions plasmidiques ont été produites pour réaliser ce travail. L'ADNc du gène de la TM a été cloné dans le plasmide pCDNA3.1 (TM-WT). Les autres constructions ont été créées par mutagenèse dirigée en utilisant le kit Quickchange XL mutagenesis (Stratagene). La seconde construction (TM- TM<sub>1-536</sub>) contient l'ADNc total de la TM avec la mutation c.1611C>A (p.Cys537X). Les constructions TM<sub>1-515</sub> (p.Gly516X), TM<sub>1-525</sub> (p.Leu526X), TM<sub>1-529</sub> (p.Val530X), TM<sub>1-537</sub> (p.His538X) correspondent à des TM mutantes ayant des domaines transmembranaires de taille variable.

### 2. Expression cellulaire des variants de la thrombomoduline

Des cellules COS-1 (ATCC, CRL-1650) ont été cultivées à 37°C sous 5% de CO<sub>2</sub> dans du milieu IMDM en présence de 10% de SVF et 1% de pénicilline et streptomycine. 5,2X10<sup>4</sup> cellules/cm<sup>2</sup> ont étéensemencées dans des plaques 6 ou 24 puits et transfectées transitoirement à l'aide du réactif XtremeGENE9 DNA (Roche molecular diagnostics) pré-incubé avec 0,2µg de plasmide pour 1X10<sup>5</sup> cellules dans du milieu IMDM sans SVF. L'analyse de l'expression des différents mutants de la TM a été réalisée 36 ou 48h après transfection.

### 3. Western-blot

Trente-six heures après transfection, les surnageants cellulaires ont été éliminés et remplacés par du milieu IMDM sans SVF puis les cellules ont été incubées une nuit à 37°C. Le lendemain, les surnageants cellulaires ont été récoltés, centrifugés et conservés à -20°C. Après deux lavages en DPBS 1X les tapis cellulaires ont été lysés par un tampon de lyse (HEPES pH 7.5 50mM, KCl 0.1M, MgCl<sub>2</sub> 2mM, Triton X-100 0.5%) contenant un inhibiteur de protéase (Complete Mini) 20 minutes dans la glace. Après centrifugation, les lysats cellulaires ont été conservés à -20°C.

Avant utilisation, les surnageants cellulaires ont été concentrés à l'aide du kit Amicon Ultra 0,5mL centrifugal filter 10K (Merck KGaA, Darmstadt, Deutschland). Après migration sur un gel SDS-PAGE 10% et transfert, la TM a été visualisée dans les lysats cellulaires et les surnageants de culture à l'aide d'un anticorps monoclonal anti-TM humaine (anti-thrombomodulin antibody [PBS-02] ab7640, Abcam) dilué au 1/3000 et un anticorps secondaire polyclonal anti-souris couplé à une peroxidase (Goat Anti-Mouse IgG (H+L)-HRP Conjugate, Hercules, CA, USA). La révélation par chimiluminescence a été réalisée à l'aide du kit Immun-Star™ (Western C™ Chemiluminescent, BioRad, Hercules, CA, Etats-Unis).

### 4. Dosage de la thrombomoduline par ELISA

La concentration de TM dans les surnageants et les lysats cellulaires a été déterminée à l'aide du Kit ELISA de chez Abcam.

### 5. Etude de l'action de protéases et d'agents chimiques sur la libération de TM

L'action de différentes protéases et agents chimiques sur la libération de TM dans le milieu cellulaire a été testée. Les cellules COS-1 ont été cultivées en plaques 24 puits puis 48h après transfections les surnageants cellulaires ont été éliminés et remplacés par exactement 250 µl de milieu IMDM sans SVF contenant soit une protéase soit un agent chimique. Les surnageants cellulaires ont été récoltés à 30, 60 et 90 minutes, centrifugés puis congelés. Les protéases testées étaient : la cathepsine G (ICN biochemicals, GmbH, Eschwege, Allemagne) et l'élastase (Sigma-Aldrich, Saint Louis, MI, USA) aux concentrations 1µg/ml, 2,5 µg/ml et 5µg/ml. Les agents chimiques (Sigma-



Aldrich) testés étaient le L-Gluthation réduit (GSH), le peroxyde d'hydrogène et la N-acétylcystéine (NAC). Un inhibiteur de métalloprotéase a aussi été utilisé : le GM6001 (Calbiochem, San Diego, CA, Etats-Unis).

#### 6. Etude de la libération de TM en présence de forces de cisaillement

L'effet des forces de cisaillement sur la libération de TM a été évalué. Les cellules COS-1 ont été cultivées en plaques 24 puits puis 48h après transfection, les cellules ont été soumises à des forces de cisaillement ( $5000s^{-1}$ ) reproduisant les conditions existantes dans les capillaires pendant différents temps (0, 2, 5 et 10 minutes). Ces forces ont été appliquées à l'aide d'un viscosimètre de type cône et plan. Les concentrations de TM dans les surnageants cellulaires ont ensuite été déterminées par ELISA.

#### 7. Marquage des tapis cellulaire par immunofluorescence

L'expression cellulaire de la TM à la surface des cellules après transfection a été vérifiée par immunofluorescence. Les cellules COS-1 ont été cultivées dans des cellules Labtek (Nalge Nunc, Naperville, USA). Quarante-huit heures après transfection, le surnageant cellulaire a été éliminé, les cellules ont été lavées avec du PBS et fixées avec du paraformaldéhyde à 4% pendant 15 min à température ambiante. Après trois lavages en PBS, les cellules ont été saturées avec du milieu IMDM contenant 0,5% de BSA pendant 30 min. Ensuite la TM présente à la surface des cellules a été marquée à l'aide d'un anticorps monoclonal de souris anti-TM (Abcam) dilué au 1/1000 dans du milieu IMDM 0,5% de BSA (1h d'incubation à température ambiante). Les membranes cellulaires ont été marquées à l'aide d'agglutinines conjuguées à l'Alexa Fluor 594® (Wheat Germ Agglutinin conjugate, Invitrogen, Carlsbad, USA) diluées ( $5\mu g/ml$ ) dans du milieu IMDM 0,5% de BSA (10 min d'incubation à température ambiante). Après 3 lavages en PBS, les cellules ont été incubées en présence d'un anticorps secondaire anti-souris conjugué à l'Alexa Fluor 488® (Invitrogen) dilué au 1/800 dans du milieu IMDM 0.5% de BSA (30 min d'incubation à température ambiante). Après trois lavages, les puits ont été séchés puis la lame a été montée. Un microscope Leica SP5 X (Leica, Wetzlar, Allemagne) a été utilisé pour réaliser les images d'immunofluorescence.

#### 8. Test de génération de thrombine

L'activité de la TM présente dans les surnageants de culture cellulaire a été évaluée en mesurant la diminution de la génération de thrombine d'un plasma pauvre en plaquettes (PPP) témoin. Les générations de thrombine ont été mesurées par la méthode CAT (Calibrated Automated Thrombography, Stago, Asnières, France). Dans une plaque de 96 puits (Greiner, Poitiers, France) ont été mélangés 80 µl de plasma pauvre en plaquettes (PPP) témoin, 10 µl d'un mélange contenant du facteur tissulaire (1pM) et des phospholipides (concentration finale 4µM, composition : 20% phosphatidylsérine, 20% phosphatidyléthanolamine, 60% phosphatidylcholine, Avanti Polar lipids, Alabaster, AL, Etats-Unis) et 10µl de surnageant cellulaire. Après ajout de 20 µl de réactif contenant du calcium et un substrat fluorescent de la thrombine (FluCa). La fluorescence émise à 460 nm (longueur d'onde d'excitation 390 nm) a été suivie au cours du temps. La courbe de génération de thrombine et le calcul des différents paramètres ont été réalisés par le logiciel Trombinoscope version 3.0.0.26 (Trombinoscope BV). Le principal paramètre utilisé dans ce travail est le potentiel endogène de thrombine (ETP). L'ETP correspond à l'aire sous la courbe de génération de thrombine et il a déjà été démontré que ce paramètre est mieux corrélé aux phénotypes cliniques de patients atteints de pathologies hémorragiques que les tests standards de coagulation (314,315).

#### 9. Analyse statistique des résultats

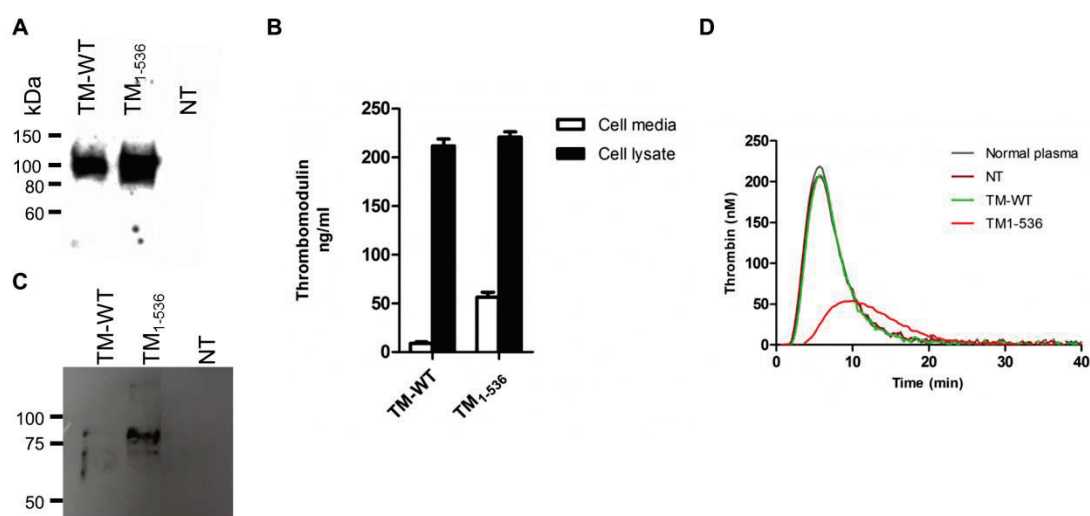
L'analyse statistique des résultats a été réalisée à l'aide du logiciel GraphPad Prism InStat 3.0 (La Jolla, CA, Etats-Unis). Les résultats sont présentés sous la forme de moyenne (m) +/- écart-type (sd). La comparaison entre les moyennes a été effectuée à l'aide du test t de Student. Une valeur de  $p < 0,05$  a été considérée comme statistiquement significative.

## 4. Résultats

### **Expression cellulaire du mutant TM1-536 (p.Cys537X) dans les cellules COS-1**

La mutation *THBD* c.1611 C>A (p.Cys537X) entraîne la synthèse d'une protéine tronquée ayant perdu à la fois les 3 derniers aa du domaine transmembranaire et le domaine intracellulaire. La TM-WT et la TM<sub>1-536</sub> ont été transitoirement transfectées dans des cellules COS-1. Quarante-huit heures

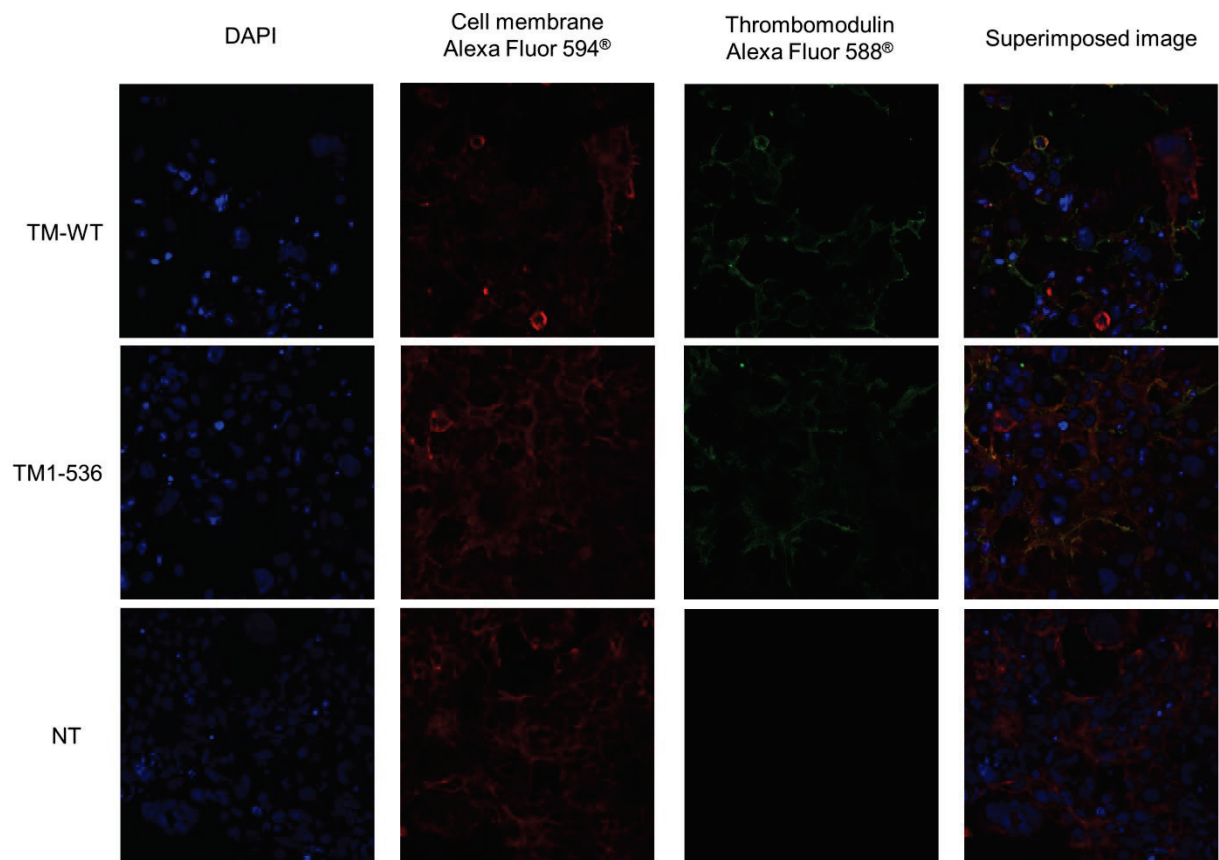
après transfection, des western-blots ont été réalisés sur les lysats cellulaires et les surnageants cellulaires. Une bande de 100 kDa correspondant à la TM a été détectée dans les lysats des cellules exprimant la TM-WT ou la TM<sub>1-536</sub> (Figure 48A). Les quantités de TM présente dans les lysats cellulaires étaient similaires (211 +/- 7 vs 220 +/- 5 ng.ml<sup>-1</sup> respectivement pour la TM-WT et la TM<sub>1-536</sub>, *p* = 0,37 ; Figure 48B). Cependant des taux de TM significativement plus élevés ont été observés dans les surnageants cellulaires des cellules exprimant la TM<sub>1-536</sub> par rapport à celles exprimant la TM-WT (56,3 +/- 5,2 vs 8, 8 +/- 1,6 ng.ml<sup>-1</sup> respectivement ; Figure 48B). Ces résultats ont été confirmés par des western-blots réalisés sur les surnageants cellulaires. Une bande unique de 75-100 kDa correspondant à la TM<sub>1-536</sub> a été observée. Aucune bande n'est présente sur le western-blot pour la TM-WT (Figure 48C). Ces résultats démontrent que la TM<sub>1-536</sub> est correctement synthétisée par les cellules COS-1 transfectées et qu'une fraction de la TM<sub>1-536</sub> est sécrétée dans le milieu extracellulaire.



**Figure 48 : Expression cellulaire des TM-WT et TM<sub>1-536</sub> (p.Cys537X).** A : western-blot en condition réductrice réalisé sur les lysats cellulaires des cellules transfectées par le plasmide codant la thrombomoduline naturelle (TM-WT), la thrombomoduline mutée c.1611C<A (p.Cys537X, TM<sub>1-536</sub>) ou des cellules non transfectées (NT). B : Dosage par ELISA de la thrombomoduline présente dans les surnageants et les lysats cellulaires. C : western-blot en condition réductrice réalisé sur les surnageants cellulaires des cellules transfectées par le plasmide codant la thrombomoduline naturelle (TM-WT), la thrombomoduline mutée c.1611C<A (p.Cys537X, TM<sub>1-536</sub>) ou des cellules non transfectées (NT). D : Mesure de l'activité anticoagulante de la thrombomoduline présente dans les surnageants de culture cellulaire. Cette activité a été évaluée en mesurant la diminution de la génération de thrombine d'un plasma pauvre en plaquettes témoin en présence de surnageant cellulaire.

L'expression à la surface des cellules COS-1 des TM-WT et TM<sub>1-536</sub> a été étudiée par immunofluorescence. Deux marquages ont été utilisés : le premier pour la TM (Alexa Fluor 488<sup>®</sup>) et le

second pour la membrane cellulaire (Alexa Fluor 594®). Les différentes images obtenues avec ces marquages sont présentées au niveau de la Figure 49. La présence d'un double marquage (zone jaune) indique l'expression des TM-WT et TM<sub>1-536</sub> à la surface des cellules COS-1 transfectées. Ce résultat confirme qu'une fraction de la TM<sub>1-536</sub> est ancrée dans la membrane plasmique tandis qu'une autre fraction de ce variant de la TM est sécrété dans le milieu extracellulaire.



**Figure 49 : Analyse par immunofluorescence de la thrombomoduline à la surface des cellules COS-1.** Les cellules COS-1 ont été transitoirement transfectées par les plasmides TM-WT et TM<sub>1-536</sub>, incubées pendant 48 h puis marquées par immunofluorescence. Les noyaux cellulaires ont été marqués par du DAPI (bleu). Les membranes plasmiques ont été marquées par une agglutinine couplée à l'Alexa Fluor 594® (rouge) et la TM par un anticorps couplé à l'Alexa Fluor 488® (vert). La superposition des différentes images montre que la TM-WT et la TM<sub>1-536</sub> sont exprimées à la surface des cellules COS-1 (jaune). NT : cellules non transfectées.

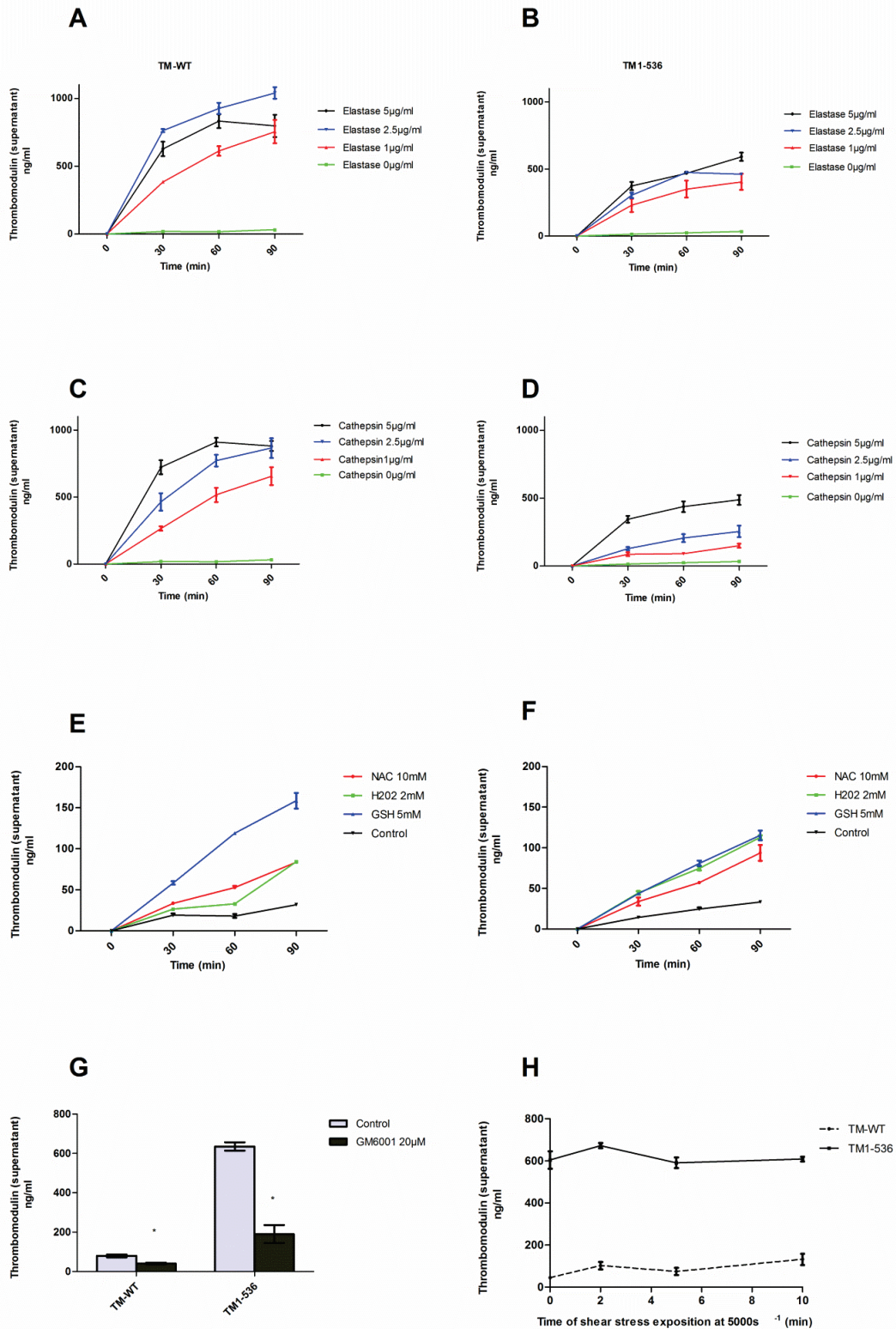
### La TM<sub>1-536</sub> sécrétée conserve son activité anticoagulante

L'activité anticoagulante de la TM sécrétée a été évaluée en mesurant la diminution de la génération de thrombine d'un plasma contrôle (ETP : 1291 nM.min) lors de l'ajout de 10 µl de surnageant de cellules exprimant soit la TM-WT soit la TM<sub>1-536</sub> (Figure 48 D). Aucune diminution

significative de la génération de thrombine n'a été observée lors de l'ajout de 10µl de surnageant de cellules non transfectées ou de cellules exprimant la TM-WT à 80µl de PPP témoin (ETP = 1269 et 1204 nM.min respectivement). A l'inverse, une diminution de 54 % de l'ETP a été observée lorsque le surnageant des cellules exprimant la TM<sub>1-536</sub> a été mélangé au PPP témoin (ETP = 595 nM.min). Ces résultats démontrent que la fraction sécrétée de la TM<sub>1-536</sub> conserve son activité anticoagulante.

### **La TM<sub>1-536</sub> n'est pas plus sensible à l'action de certaines protéases et des agents réducteurs.**

Une faible concentration de sTM est présente physiologiquement dans le plasma. Cette fraction est issue du clivage protéolytique ou chimique de la TM présente à la surface des cellules endothéliales. Nous avons émis l'hypothèse que la TM<sub>1-536</sub>, du fait de la présence d'un domaine transmembranaire raccourci, pourrait avoir une plus grande sensibilité au clivage par les protéases ou par les agents chimiques. Quarante-huit heures après transfection, le milieu cellulaire a été éliminé et remplacé par un milieu sans SVF contenant différentes concentrations de protéases (cathepsine G et élastase : 1, 2,5 et 5 µg.ml<sup>-1</sup>) ou d'agents chimiques réducteurs (NAC 10 mM, H<sub>2</sub>O<sub>2</sub> 2mM et GSH 5 mM). La libération de TM dans les surnageants cellulaires a été mesurée après 30, 60 et 90 minutes d'incubation à 37°C. La TM-WT et la TM<sub>1-536</sub> sont rapidement libérées dans le surnageant cellulaire en présence de cathepsine G et d'élastase. Un relargage important de ces deux protéines est détecté dès 30 minutes d'incubation en présence de protéase (Figure 50 A à D). La majorité de la TM est retrouvée dans les surnageants cellulaires après 1 h d'exposition aux protéases. La concentration dans les milieux cellulaires de TM<sub>1-536</sub> est inférieure à celle de TM-WT car une partie de la TM<sub>1-536</sub> a été sécrétée spontanément dans le surnageant cellulaire avant l'incubation avec les protéases, et cette fraction a été éliminée lors du changement de milieu. L'incubation des cellules COS-1 en présence d'agents réducteurs entraîne la libération de TM-WT et de TM<sub>1-536</sub> de façon équivalente (Figure 50 E et F).



**Figure 50 : libération de thrombomoduline dans les surnageants cellulaires sous différentes conditions.** Les concentrations de thrombomoduline dans les surnageants cellulaires ont été mesurées par ELISA. A à D : impact de la présence de protéases sur la libération de thrombomoduline. Les cellules transfectées ont été incubées différents temps en présence de cathepsine G (1, 2,5, 5  $\mu\text{g}\cdot\text{ml}^{-1}$ ) ou d'élastase (1, 2,5, 5  $\mu\text{g}\cdot\text{ml}^{-1}$ ). E et F : impact de la présence d'agents chimiques réducteurs sur la libération de thrombomoduline. Les cellules transfectées ont été incubées différents temps en présence de peroxyde d'hydrogène ( $\text{H}_2\text{O}_2$  2mM), de N-acétyl-L-cystéine (NAC, 10 mM) et de glutathion réduit (GSH, 5mM). G : Impact de la présence de GM60001, un inhibiteur à large spectre de métalloprotéases, sur la libération de TM-WT



et TM<sub>1-536</sub>. Après transfection, les cellules COS-1 ont été incubées 48h en présence de GM6001 (20µM). Les résultats sont présentés sous la forme d'une moyenne +/- l'écart type (n >3). \**p* < 0,05. H: libération de thrombomoduline après exposition des cellules COS-1 à des forces de cisaillement. Quarante-huit heures après transfection, les cellules COS-1 ont été exposées à des forces de cisaillement (5000 s<sup>-1</sup>) à l'aide d'un viscosimètre de type « cône et plan » pendant 0, 2, 5 et 10 minutes. La concentration de thrombomoduline dans les surnageants cellulaires a été ensuite mesurée pour en évaluer la libération.

### **L'inhibiteur de métalloprotéases GM6001 diminue de manière importante la libération de TM<sub>1-536</sub> dans les surnageants de culture.**

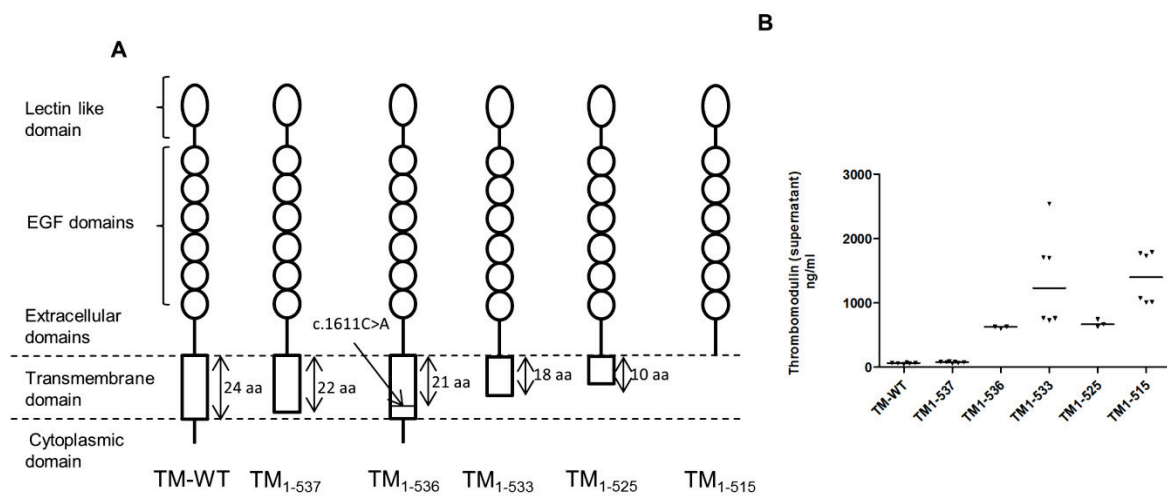
Il a été décrit que plusieurs métalloprotéases sont impliquées dans le clivage protéolytique des domaines extracellulaires des protéines transmembranaires. Afin d'évaluer le rôle des métalloprotéases dans la libération de TM<sub>1-536</sub>, des cellules transfectées ont été cultivées en présence de GM6001 (20µM), un inhibiteur à large spectre de métalloprotéases (ADAMs et métalloprotéases matricielles ; MMP). Quarante-huit heures après transfection, la libération de TM dans les surnageants cellulaires a été mesurée. Une diminution de 75 % de la libération de TM<sub>1-536</sub> a été observée dans le cas des cellules cultivées en présence de GM6001 (Figure 50 G). Cependant la concentration de sTM reste significativement plus élevée dans les surnageants de cellule exprimant la TM<sub>1-536</sub> que celles exprimant la TM-WT. Ainsi, il semblerait que l'une des raisons de l'importante libération de la TM<sub>1-536</sub> soit liée à l'action de métalloprotéases.

### **Les forces de cisaillement n'entraînent pas de libération de la TM<sub>1-536</sub> dans le milieu extracellulaire.**

La TM est très fortement exprimée au niveau de l'endothélium de vaisseaux sanguins soumis à de très grandes forces de cisaillement tels que les capillaires sanguins (2000 à 5000 s<sup>-1</sup>). Nous avons émis l'hypothèse que la présence d'un domaine transmembranaire de taille réduite pourrait être responsable d'une anomalie d'ancrage de la protéine dans la membrane cytoplasmique et donc que la TM<sub>1-536</sub> soit libérée dans la circulation sanguine sous l'action des forces de cisaillement. Afin de vérifier notre hypothèse, 48h après transfection, des cellules COS-1 ont été soumises à des forces de cisaillement (5000s<sup>-1</sup>) pendant différents temps (0, 2, 5, 10 minutes) puis nous avons mesuré par ELISA les concentrations de sTM dans les surnageants cellulaires. Aucune augmentation significative de la concentration de TM n'a été mise en évidence après exposition des cellules aux forces de cisaillement (*p* = 0,229 pour la TM-WT et *p* = 0,078 pour la TM<sub>1-536</sub>; Figure 50 H).

## La libération spontanée de TM dans les surnageants cellulaires est dépendante de la taille du domaine intramembranaire.

Dans l'objectif d'étudier la relation existante entre la taille du domaine transmembranaire résiduel et la libération spontanée de TM dans les surnageants cellulaires, 5 TM mutantes ayant des domaines transmembranaires de taille variable (allant de l'absence totale de domaine transmembranaire pour la TM<sub>1-515</sub> à un domaine transmembranaire complet pour la TM-WT) ont été exprimé dans des cellules COS-1 (Figure 51 A). Quarante-huit heures après transfection, les concentrations de sTM dans les surnageants cellulaires ont été déterminées par ELISA (Figure 51 B). Des concentrations plus élevées de sTM ont été observées pour les variants TM<sub>1-536</sub> (618 +/- 8 ng.ml<sup>-1</sup>), TM<sub>1-533</sub> (1366 +/- 303 ng.ml<sup>-1</sup>), TM<sub>1-525</sub> (688 +/- 33 ng.ml<sup>-1</sup>) et TM<sub>1-515</sub> (1395 +/- 164 ng.ml<sup>-1</sup>) en comparaison aux taux de sTM observés pour les variants TM-WT (62 +/- 4 ng.ml<sup>-1</sup>,  $p < 0,05$ ). Aucune différence n'a été trouvée entre les variants TM-WT et TM<sub>1-537</sub> (62 +/- 4 vs 75 +/- 4 ng.ml<sup>-1</sup>,  $p = 0.0612$ ).



**Figure 51 : Relation entre la taille du domaine transmembranaire et la libération de thrombomoduline dans les surnageants cellulaires.** A : représentation schématisques des différentes thrombomodulines mutantes étudiées. B : libération spontanée de thrombomoduline dans les surnageants de cellules exprimant les différents mutants testés. aa : acides aminés.

Ces résultats montrent que la taille du domaine transmembranaire est un élément important de l'ancrage de la TM dans la membrane cytoplasmique. La réduction de la taille du domaine transmembranaire entraîne des anomalies d'ancrage des protéines se traduisant par une libération



spontanée de la TM dans le milieu extracellulaire de manière indépendante de tous les autres paramètres.

## 5. Discussion

La plupart des mutations du gène *THBD* ont été décrites chez des patients ayant des maladies thrombo-emboliques veineuses ou des syndromes hémolytiques et urémiques atypiques (316–318). Récemment, deux équipes indépendantes ont décrit une nouvelle maladie hémorragique de transmission autosomique dominante chez des patients porteurs de la mutation *THBD* c.1611C>A (p.Cys536X) (130,131). Cette mutation est associée à la présence d'une concentration très élevée de sTM dans le plasma des patients. Il a été démontré que cette concentration importante de sTM était responsable d'une augmentation de la génération de PCa entraînant la dégradation très rapide des FVIIIa et FVa ainsi qu'une diminution de la génération de thrombine chez les propositus. L'objectif de cette étude a été de déterminer les mécanismes moléculaires impliqués dans la libération dans la circulation de la TM chez les patients porteurs de la mutation *THBD* c.1611C>A.

La mutation *THBD* c.1611C>A entraîne la production d'un TM tronquée dans son domaine transmembranaire (perte des 3 derniers acides aminés du domaine transmembranaire ainsi que la totalité du domaine intracellulaire) (130,131). Dans un premier temps, nous avons voulu savoir si la TM<sub>1-536</sub> était correctement exprimée par les cellules endothéliales. Pour cela, nous avons développé un modèle d'expression cellulaire de variant de la TM dans des cellules COS-1. Les résultats des western-blot et des dosages ELISA réalisés dans les lysats de cellules transfectées par des vecteurs plasmidiques codant la TM<sub>1-536</sub> et la TM-WT ont montré que ces deux protéines étaient exprimées de manière équivalente par les cellules COS-1. De plus, nous avons démontré par marquage immunofluorescent que les TM<sub>1-536</sub> et TM-WT étaient ancrées à la surface des cellules transfectées. Cependant, un taux significativement plus élevé de sTM a été observé dans les surnageants de cellules exprimant la TM<sub>1-536</sub> par rapport à celle exprimant la TM-WT. Ces résultats suggèrent que la TM<sub>1-536</sub> est en partie sécrétée dans le domaine extracellulaire. Ces données *in vitro* sont en accord avec les observations *in vivo* montrant à la fois la présence d'une quantité importante de sTM dans le plasma des patients porteurs de la mutation c.1611C>A ainsi qu'une diminution de l'expression de la TM à la

surface des cellules endothéliales (observation réalisée sur une biopsie de tissus ovariens) (131). Dans cette étude, nous avons aussi démontré que la TM<sub>1-536</sub> libérée dans le milieu extracellulaire conservait son activité anticoagulante. Ce résultat est conforté par les phénotypes hémorragiques présents chez les patients ainsi que par les données des études des relations structure/activité de la TM qui ont montré que l'activité anticoagulante de la TM est portée par les domaines EGF 4, 5 et 6 (319).

En plus de son activité anticoagulante liée à la génération de PCa, le complexe TM/thrombine a aussi une activité antifibrinolytique liée à l'activation du TAFI (293). La génération de TAFIa par le complexe TM/thrombine nécessite la présence des domaines EGF 3 à 6 qui sont intacts sur la TM<sub>1-536</sub> (277). Il a été montré que la sTM présente dans la circulation des patients porteurs de la mutation c.1611C>A a une activité antifibrinolytique intacte (133). Cependant, il est probable que la très faible génération de thrombine observée chez les *propositus* limite les capacités du complexe TM/thrombine à activer le TAFI et donc que cette activité antifibrinolytique ne soit pas suffisante pour contrecarrer l'activité anticoagulante de la sTM.

Dans les conditions physiologiques, la sTM est présente en faible concentration dans la circulation et est issue du relargage de la TM exprimée à la surface des cellules endothéliales. La libération d'ectodomaine (relargage dans le milieu extracellulaire de domaines extracellulaires de protéines exprimées à la surface des cellules) est un mécanisme utilisé par les cellules pour réguler sélectivement l'expression de protéines de surface (320). Dans le cas de la TM, les enzymes impliquées dans ce phénomène (les *sheddas*) sont inconnues mais des données *in vitro* semblent indiquer que plusieurs métalloprotéases appartenant aux familles ADAMs et MMP sont impliquées (283). Nous avons émis l'hypothèse que les taux très élevés de sTM présents dans les sérums de patients porteurs de la mutation c.1611C>A étaient liés à une plus grande sensibilité de la TM<sub>1-536</sub> aux *sheddas*. L'utilisation d'inhibiteurs de métalloprotéases à large spectre entraîne une diminution significative de la libération de sTM par les cellules exprimant la TM<sub>1-536</sub>. Ce résultat met en évidence le rôle des métalloprotéases sur la libération de TM<sub>1-536</sub>. Il est possible que l'absence d'une partie du domaine transmembranaire et du domaine intracellulaire rende la TM<sub>1-536</sub> plus sensible à l'action des métalloprotéases que la TM-WT. Il a été récemment montré que la TM était un substrat spécifique de

la rhomboïde 2 (RHBDL2), une sérine protéase présente à la surface des cellules de mammifères (282,321). Le clivage de la TM<sub>1-536</sub> par la RHBDL2 n'a pas été étudié car le domaine intracellulaire est indispensable à la reconnaissance de son substrat par la RHBDL2 (282). Enfin, plusieurs processus d'origine inflammatoire sont associés à une augmentation modérée, mais significative, du taux de sTM dans le plasma (322,323) du fait à la fois de la protéolyse de la TM membranaire par différentes protéases libérées par les leucocytes (élastase et cathepsine G) et de l'action d'agents chimiques (281,283). Dans cette étude aucune différence n'a été observée entre les libérations de TM-WT et de TM<sub>1-536</sub> après incubation des cellules en présence d'élastase, de cathepsine G et de différents agents chimiques.

Le rôle potentiel des forces de cisaillements sur la libération de TM<sub>1-536</sub> a aussi été exploré dans ce travail. La TM est fortement exprimée dans les vaisseaux sanguins de faible calibre comme les capillaires où existent des forces de cisaillement très importantes (324). L'introduction d'un groupe chargé négativement à l'extrémité C-terminale de la TM dans la bicouche lipidique pourrait déstabiliser l'ancrage de la TM<sub>1-536</sub> dans la membrane plasmique. Cependant nos résultats ont montré que l'application de forces de cisaillement aux cultures cellulaires ne n'augmentait pas la libération de TM<sub>1-536</sub> dans le milieu extracellulaire.

Dans ce travail nous avons aussi exploré la relation entre la taille du domaine transmembranaire résiduel et la libération de TM dans le milieu extracellulaire. Des taux élevés de sTM dans les surnageants de culture ont été observés pour tous les mutants ayant un domaine transmembranaire diminué d'au moins 3 acides aminés. Ces résultats démontrent que la libération de la sTM est dépendante de la présence ou non d'un domaine transmembranaire entier. Il est possible que lors de la biosynthèse de la TM dans la cellule endothéliale, la présence d'une protéine tronquée au niveau du domaine transmembranaire soit responsable d'une anomalie de localisation cellulaire et qu'une partie de cette protéine anormale soit sécrétée au lieu d'être ancrée dans la membrane.

La principale limitation de cette étude est l'utilisation de cellules COS-1 qui n'expriment pas de manière constitutive la TM. Ces résultats préliminaires devront être confirmés sur une lignée de cellules endothéliales telle que les HUVECs.

En conclusion, dans ce travail nous avons étudié les conséquences moléculaires de la mutation gain de fonction *THBD* c.1611C>A (p.Cys573X). Nos résultats ont montré que les mécanismes responsables de la libération de la TM<sub>1-536</sub> dans le milieu extracellulaire étaient complexes et impliquaient au moins deux processus liés à la présence d'un domaine transmembranaire tronqué: une plus grande sensibilité aux métalloprotéases ainsi qu'une anomalie de biosynthèse.

## Conclusions et perspectives

Le développement de nouvelles techniques de biologie moléculaire depuis les années 1990 a permis de détecter des anomalies génétiques chez un nombre toujours plus grand de patients atteints de pathologies hémorragiques héréditaires. Dans l'exemple de l'HA, la première analyse complète des régions codantes du gène *F8* a été publiée en 1991 (325). Cependant, aucune anomalie génétique n'était trouvée dans les régions codantes du gène chez plus de la moitié des hémophiles A sévères (326). Quelques années plus tard, en 1993, le phénomène d'Inv22 et le développement d'une méthode de détection de ce phénomène par Southern blot ont permis de montrer que cette anomalie était présente chez 45% des HA sévères (56). Par la suite, des méthodes plus simples de détection de l'Inv22, basées sur le principe de la PCR, ont été publiées (327). En 2001, le phénomène Inv1 est découvert et identifié chez 2% des HA sévères (67). A la même période, l'amélioration des méthodes de recherche de mutations dans les régions codantes avec notamment l'apparition de l'analyse par chromatographie liquide à haute performance en condition dénaturante a entraîné une augmentation de la sensibilité de l'analyse du gène *F8*. Cependant, des anomalies étaient identifiées que dans 70 à 80% des cas d'HA (328). Certaines anomalies restaient difficiles à identifier et l'apparition de la MLPA en 2008 a facilité la détection des duplications et des délétions de grande taille (329). Enfin, plus récemment des mutations introniques ont été identifiées comme responsables d'HA (124). Ainsi, l'amélioration continue des techniques de biologie moléculaire a permis d'augmenter du taux de détection de mutation au cours du temps chez les hémophiles A. Un nombre de plus en plus important de familles ont alors pu bénéficier d'un conseil génétique plus précis.

Actuellement, les performances des méthodes de séquençage à haut débit autorisent l'analyse d'un nombre très important de patients et permettent de ne plus se concentrer uniquement sur les patients présentant les phénotypes les plus sévères. En conséquence, de nouveaux variants localisés dans les régions codantes ou au niveau des jonctions intron/exon sont régulièrement décrits. En parallèle, les techniques d'étude globale du génome telle que l'analyse chromosomique sur puce à

ADN (CMA) permettent de mettre en évidence l'existence d'anomalies génétiques de grande taille ne se limitant pas à un seul gène. L'impact clinique de ces différents types d'anomalies génétiques est souvent difficile à évaluer pour le généticien. Cela rend délicat le conseil génétique donné aux individus porteurs de ces variants.

La première partie de ce travail de thèse a été centrée sur l'étude des corrélations génotype / phénotype chez les patients hémophiles porteurs de réarrangements génomiques de grande taille. Dans une première étude, nous avons étudié par CMA 6 patients porteurs d'une délétion totale du gène *F9* dont 5 d'entre eux présentaient des déficits mentaux plus ou moins importants. Les analyses par CMA ont montré l'existence chez tous ces patients de délétions de très grande taille englobant de nombreux gènes autres que le gène *F9*. A l'aide de nos résultats et des données de la littérature, nous avons identifié le gène *SOX3* comme candidat probable au déficit mental observé chez les patients. Nous avons réalisé une étude similaire chez des patients hémophiles A ayant des profils anormaux d'Inv 22 et d'Inv1. Les études de CMA réalisées chez ces patients ont montré l'existence de réarrangements très complexes de la région Xq28 impliquant de nombreux gènes voisins du gène *F8*. Parmi ces gènes, plusieurs sont associés à des pathologies graves telles que des maladies cardiovasculaires ou des déficits mentaux. Ces deux études ont un impact direct sur la qualité du conseil génétique donné aux patients et à leur famille en permettant une description plus précise du type de pathologies pouvant être transmises à leur descendance (exemple : hémophilie B sévère et déficit mental). La réalisation de ce type d'étude doit maintenant être élargie aux patients porteurs de délétion des premiers ou derniers exons des gènes *F8* ou *F9* afin d'identifier si des gènes voisins sont aussi concernés par l'anomalie génétique. Ce type d'étude souligne aussi l'importance de la coopération entre les médecins et les biologistes moléculaires. En effet une description précise des phénotypes cliniques, ne se limitant pas aux symptômes hémorragiques, peut aider et encourager le biologiste moléculaire à rechercher d'autres anomalies génétiques.

La deuxième partie de ce travail a été focalisée sur les variations de signification indéterminée localisées au niveau des sites d'épissage du gène *F8*. La détermination du caractère pathogène de ce type d'anomalie reste un défi pour les biologistes moléculaires. L'analyse de l'ARNm est une étape

indispensable pour établir avec certitude l'impact sur l'épissage de ce type de variant. En l'absence de prélèvement d'ARNm, l'utilisation conjointe d'algorithmes informatiques et de tests *in vitro* de type minigène semble être une approche efficace pour évaluer l'effet délétère de ces variants. Ainsi, le développement de ce type d'approche nous a permis de démontrer la pathogénicité de 21 variants d'épissage du gène *F8* identifiés dans notre cohorte d'hémophiles A. Le développement de méthodes *in vitro* d'étude de variant (variant faux-sens ou variant d'épissage) a plusieurs avantages. Ces méthodes permettent de fournir des arguments forts en faveur du caractère délétère d'une mutation et ainsi d'éviter d'identifier à tort des polymorphismes rares comme des variants pathogènes. Ce type d'erreur peut être responsable de la dispensation d'un conseil génétique erroné et donc avoir des conséquences sur un éventuel diagnostic prénatal comme cela a été rapporté en 2009 par Schneppenheim *et al.* (330). Enfin, la possibilité de séquencer le gène *F8* entier chez des patients chez lesquels aucune anomalie n'est encore caractérisée va entraîner l'identification de nombreux variants de signification indéterminés dans les régions introniques. Des méthodes bio-informatiques pour éliminer les variants sans intérêts et des études *in vitro* deviennent alors indispensables pour avoir une interprétation correcte de l'effet de ces anomalies. Ce point a été particulièrement développé dans la troisième partie de cette thèse.

Aucune anomalie génétique n'est détectée dans un nombre non négligeable d'individus atteints de pathologies hémorragiques congénitales. La dernière partie de ce travail a été consacrée à la recherche de nouveaux mécanismes moléculaires permettant d'expliquer les phénotypes cliniques de ces patients. Dans une première étude, nous avons identifié une délétion intronique profonde localisée dans l'intron 13 du gène *F8* et présente chez 6,1% des hémophiles A mineurs issus de notre cohorte. Nos modèle *in vitro* nous ont permis de montrer que cette anomalie génétique était probablement responsable d'une dérégulation de l'hnRNP C ce qui a pour conséquence l'insertion d'une séquence *AluY* dans les transcrits du gène *F8*. Du fait de la haute prévalence de cette anomalie dans la population française d'hémophiles A, cette délétion doit être spécifiquement recherchée chez tous les patients hémophiles A mineurs chez lesquels aucune anomalie génétique n'est mise en évidence lors de l'analyse moléculaire de routine du gène *F8*. Cependant, des recherches de cette anomalie dans

d'autres populations d'hémophiles A sont encore nécessaires afin de pouvoir statuer sur le caractère récurrent ou fondateur de cette anomalie.

Finalement, dans une seconde étude, nous avons analysé les conséquences de la présence de la mutation *THBD* c.1611C>A sur le métabolisme de la thrombomoduline. Cette anomalie a été identifiée lors d'études d'exome réalisées chez des patients ayant un phénotype hémorragique important et une concentration plasmatique en thrombomoduline soluble (sTM) très importante. A l'aide de modèles *in vitro*, nous avons démontré que la mutation *THBD* c.1611C>A était responsable de l'augmentation de la sTM circulante et ceci du fait de mécanismes moléculaires complexes non entièrement élucidés liés à la fois à l'action de métalloprotéases de type ADAM et MMP et à une anomalie de biosynthèse de la TM mutée. Nos résultats ont permis d'apporter de nouvelles données sur le métabolisme de la thrombomoduline et notamment sur le rôle des métalloprotéases sur la libération de sTM. Cependant des études supplémentaires sont nécessaires pour identifier avec précision les protéases impliquées.

En conclusion, ces travaux ont permis de détecter et de mieux appréhender certains mécanismes physiopathologiques des anomalies génétiques retrouvées dans des pathologies rares hémorragiques. Les approches *in vitro* de caractérisation des variants d'épissage et des variants faux-sens du gène *F8* développées dans ce travail peuvent être maintenant utilisées dans une activité de diagnostic à visée de conseil génétique. Néanmoins, de nouvelles méthodes d'études des régions introniques doivent encore être développées au laboratoire afin de réduire le nombre de patients sans anomalie génétique détectée (dans les cas de l'hémophilie A, encore 5% des hémophiles A non sévères et 2% des hémophiles A sévères sont encore sans mutation connue).



## Annexes

### **1. Article en revision (Américan Journal of Human Genetic): Reccurent *F8* intronic deletion in mild haemophilia A causes Alu exonization due to the decrease of hnRNP C binding.**

Oct 03, 2017

Dear Dr. Jourdy,

Thank you for sending your manuscript titled "Reccurent F8 intronic deletion found in mild hemophilia A causes Alu exonization due to the decrease of hnRNP C binding" (AJHG-D-17-00573) to AJHG The American Journal of Human Genetics. The reviewers commented favorably on your paper, but have raised some points to address. Please revise your manuscript by carefully addressing the comments raised by the reviewers found at the end of this email.

The reviewers have a few major criticisms that we agree should be addressed in the revision. Reviewer 1 commented on the need to evaluate the frequency of the mutation in publicly available databases. We agree that placing the mutation, likely a founder mutation, in context of other populations will strengthen the manuscript. Additionally, the reviewer commented on the lack of experimental work that explicitly demonstrates reduced hnRNP C binding in the mutant transcript. While this is the likely mechanism at work, we agree that this has not actually been evaluated. This should either be addressed experimentally or the writing (and title) should be modified to emphasize this is a hypothesis and not a result. Review 2 commented on the nature of the cell lines used in the functional characterization of the mutations. We agree with the reviewer's recommendation that wording should be included that clearly outline the limits of this approach rather than repeating the work in additional cell lines.

## Bibliographie

1. Santagostino E, Fasulo MR. Hemophilia a and hemophilia B: different types of diseases? *Semin Thromb Hemost.* 2013 Oct;39(7):697–701.
2. White GC 2nd, Rosendaal F, Aledort LM, Lusher JM, Rothschild C, Ingerslev J, et al. Definitions in hemophilia. Recommendation of the scientific subcommittee on factor VIII and factor IX of the scientific and standardization committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost.* 2001 Mar;85(3):560.
3. Mannucci PM, Tuddenham EG. The hemophilias--from royal genes to gene therapy. *N Engl J Med.* 2001 Jun 7;344(23):1773–9.
4. Berntorp E, Shapiro AD. Modern haemophilia care. *Lancet.* 2012 Apr 14;379(9824):1447–56.
5. Duncan EM, Rodgers SE, McRae SJ. Diagnostic testing for mild hemophilia a in patients with discrepant one-stage, two-stage, and chromogenic factor VIII:C assays. *Semin Thromb Hemost.* 2013 Apr;39(3):272–82.
6. Peyvandi F, Garagiola I, Young G. The past and future of haemophilia: diagnosis, treatments, and its complications. *Lancet Lond Engl.* 2016 Jul 9;388(10040):187–97.
7. Manco-Johnson MJ, Abshire TC, Shapiro AD, Riske B, Hacker MR, Kilcoyne R, et al. Prophylaxis versus episodic treatment to prevent joint disease in boys with severe hemophilia. *N Engl J Med.* 2007 Aug 9;357(6):535–44.
8. Peyvandi F, Garagiola I, Biguzzi E. Advances in the treatment of bleeding disorders. *J Thromb Haemost JTH.* 2016 Nov;14(11):2095–106.
9. Mahlangu J, Kuliczowski K, Karim FA, Stasyshyn O, Kosinova MV, Lepatan LM, et al. Efficacy and safety of rVIII-SingleChain: results of a phase 1/3 multicenter clinical trial in severe hemophilia A. *Blood.* 2016 Aug 4;128(5):630–7.
10. Mannucci PM. Back to the future: a recent history of haemophilia treatment. *Haemoph Off J World Fed Hemoph.* 2008 Jul;14 Suppl 3:10–8.
11. DiMichele D. Inhibitor development in haemophilia B: an orphan disease in need of attention. *Br J Haematol.* 2007 Aug;138(3):305–15.
12. Astermark J. Prevention and prediction of inhibitor risk. *Haemoph Off J World Fed Hemoph.* 2012 Jul;18 Suppl 4:38–42.
13. Giannelli F, Choo KH, Rees DJ, Boyd Y, Rizza CR, Brownlee GG. Gene deletions in patients with haemophilia B and anti-factor IX antibodies. *Nature.* 1983 May 12;303(5913):181–2.
14. Radic CP, Rossetti LC, Abelleiro MM, Candela M, Pérez Bianco R, de Tezanos Pinto M, et al. Assessment of the F9 genotype-specific FIX inhibitor risks and characterisation of 10 novel severe F9 defects in the first molecular series of Argentinian patients with haemophilia B. *Thromb Haemost.* 2013 Jan;109(1):24–33.
15. Miller CH, Benson J, Ellingsen D, Driggers J, Payne A, Kelly FM, et al. F8 and F9 mutations in US haemophilia patients: correlation with history of inhibitor and race/ethnicity. *Haemoph Off J World Fed Hemoph.* 2012 May;18(3):375–82.

16. Warrier I. Management of haemophilia B patients with inhibitors and anaphylaxis. *Haemoph Off J World Fed Hemoph.* 1998 Jul;4(4):574–6.
17. Franchini M, Coppola A, Tagliaferri A, Lippi G. FEIBA versus NovoSeven in hemophilia patients with inhibitors. *Semin Thromb Hemost.* 2013 Oct;39(7):772–8.
18. Berntorp E, Collins P, D’oiron R, Ewing N, Gringeri A, Négrier C, et al. Identifying non-responsive bleeding episodes in patients with haemophilia and inhibitors: a consensus definition. *Haemophilia.* 2011 Jan 1;17(1):e202–10.
19. Mannucci PM, Ruggeri ZM, Pareti FI, Capitanio A. 1-Deamino-8-d-arginine vasopressin: a new pharmacological approach to the management of haemophilia and von Willebrands’ diseases. *Lancet Lond Engl.* 1977 Apr 23;1(8017):869–72.
20. Uchida N, Sambe T, Yoneyama K, Fukazawa N, Kawanishi T, Kobayashi S, et al. A first-in-human phase 1 study of ACE910, a novel factor VIII-mimetic bispecific antibody, in healthy subjects. *Blood.* 2016 Mar 31;127(13):1633–41.
21. Shima M, Hanabusa H, Taki M, Matsushita T, Sato T, Fukutake K, et al. Factor VIII-Mimetic Function of Humanized Bispecific Antibody in Hemophilia A. *N Engl J Med.* 2016 May 26;374(21):2044–53.
22. Chowdary P, Lethagen S, Friedrich U, Brand B, Hay C, Abdul Karim F, et al. Safety and pharmacokinetics of anti-TFPI antibody (concizumab) in healthy volunteers and patients with hemophilia: a randomized first human dose trial. *J Thromb Haemost JTH.* 2015 May;13(5):743–54.
23. Sehgal A, Barros S, Ivanciu L, Cooley B, Qin J, Racie T, et al. An RNAi therapeutic targeting antithrombin to rebalance the coagulation system and promote hemostasis in hemophilia. *Nat Med.* 2015 May;21(5):492–7.
24. Nathwani AC, Tuddenham EGD, Rangarajan S, Rosales C, McIntosh J, Linch DC, et al. Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med.* 2011 Dec 22;365(25):2357–65.
25. Nathwani AC, Davidoff AM, Tuddenham EGD. Gene Therapy for Hemophilia. *Hematol Oncol Clin North Am.* 2017 Oct;31(5):853–68.
26. Peyvandi F, Palla R, Menegatti M, Mannucci PM. Introduction. Rare bleeding disorders: general aspects of clinical features, diagnosis, and management. *Semin Thromb Hemost.* 2009 Jun;35(4):349–55.
27. Mannucci PM, Duga S, Peyvandi F. Recessively inherited coagulation disorders. *Blood.* 2004 Sep 1;104(5):1243–52.
28. Palla R, Peyvandi F, Shapiro AD. Rare bleeding disorders: diagnosis and treatment. *Blood.* 2015 Mar 26;125(13):2052–61.
29. James AH. More than menorrhagia: a review of the obstetric and gynaecological manifestations of bleeding disorders. *Haemoph Off J World Fed Hemoph.* 2005 Jul;11(4):295–307.
30. Kadir RA, Economides DL, Sabin CA, Owens D, Lee CA. Frequency of inherited bleeding disorders in women with menorrhagia. *Lancet Lond Engl.* 1998 Feb 14;351(9101):485–9.

31. Peyvandi F, Di Michele D, Bolton-Maggs PHB, Lee CA, Tripodi A, Srivastava A, et al. Classification of rare bleeding disorders (RBDs) based on the association between coagulant factor activity and clinical bleeding severity. *J Thromb Haemost JTH*. 2012 Sep;10(9):1938–43.
32. Mumford AD, Ackroyd S, Alikhan R, Bowles L, Chowdary P, Grainger J, et al. Guideline for the diagnosis and management of the rare coagulation disorders: a United Kingdom Haemophilia Centre Doctors' Organization guideline on behalf of the British Committee for Standards in Haematology. *Br J Haematol*. 2014 Nov;167(3):304–26.
33. Lancellotti S, Basso M, De Cristofaro R. Congenital prothrombin deficiency: an update. *Semin Thromb Hemost*. 2013 Sep;39(6):596–606.
34. Mariani G, Bernardi F. Factor VII Deficiency. *Semin Thromb Hemost*. 2009 Jun;35(4):400–6.
35. Menegatti M, Peyvandi F. Factor X deficiency. *Semin Thromb Hemost*. 2009 Jun;35(4):407–15.
36. Duga S, Salomon O. Congenital factor XI deficiency: an update. *Semin Thromb Hemost*. 2013 Sep;39(6):621–31.
37. Schroeder V, Kohler HP. Factor XIII deficiency: an update. *Semin Thromb Hemost*. 2013 Sep;39(6):632–41.
38. Gouw SC, van den Berg HM, Oldenburg J, Astermark J, de Groot PG, Margaglione M, et al. F8 gene mutation type and inhibitor development in patients with severe hemophilia A: systematic review and meta-analysis. *Blood*. 2012 Mar 22;119(12):2922–34.
39. Gitschier J, Wood WI, Goralka TM, Wion KL, Chen EY, Eaton DH, et al. Characterization of the human factor VIII gene. *Nature*. 1984 Nov 22;312(5992):326–30.
40. Wood WI, Capon DJ, Simonsen CC, Eaton DL, Gitschier J, Keyt B, et al. Expression of active human factor VIII from recombinant DNA clones. *Nature*. 1984 Nov 22;312(5992):330–7.
41. Levinson B, Kenwrick S, Lakich D, Hammonds G, Gitschier J. A transcribed gene in an intron of the human factor VIII gene. *Genomics*. 1990 May;7(1):1–11.
42. Levinson B, Kenwrick S, Gamel P, Fisher K, Gitschier J. Evidence for a third transcript from the human factor VIII gene. *Genomics*. 1992 Nov;14(3):585–9.
43. Levinson B, Bermingham JR, Metzberg A, Kenwrick S, Chapman V, Gitschier J. Sequence of the human factor VIII-associated gene is conserved in mouse. *Genomics*. 1992 Jul;13(3):862–5.
44. Peters MF, Ross CA. Isolation of a 40-kDa Huntingtin-associated protein. *J Biol Chem*. 2001 Feb 2;276(5):3188–94.
45. Valleix S, Jeanny JC, Elsevier S, Joshi RL, Fayet P, Bucchini D, et al. Expression of human F8B, a gene nested within the coagulation factor VIII gene, produces multiple eye defects and developmental alterations in chimeric and transgenic mice. *Hum Mol Genet*. 1999 Jul;8(7):1291–301.
46. Oldenburg J, Pezeshkpoor B, Pavlova A. Historical review on genetic analysis in hemophilia A. *Semin Thromb Hemost*. 2014 Nov;40(8):895–902.
47. Kasper CK, Lin JC. Prevalence of sporadic and familial haemophilia. *Haemoph Off J World Fed Hemoph*. 2007 Jan;13(1):90–2.
48. Strauss HS. The perpetuation of hemophilia by mutation. *Pediatrics*. 1967 Feb;39(2):186–93.

49. Tuddenham EG. Flip tip inversion and haemophilia A. *Lancet Lond Engl.* 1994 Feb 5;343(8893):307–8.
50. Becker J, Schwaab R, Möller-Taube A, Schwaab U, Schmidt W, Brackmann HH, et al. Characterization of the factor VIII defect in 147 patients with sporadic hemophilia A: family studies indicate a mutation type-dependent sex ratio of mutation frequencies. *Am J Hum Genet.* 1996 Apr;58(4):657–70.
51. Leuer M, Oldenburg J, Lavergne JM, Ludwig M, Fregin A, Eigel A, et al. Somatic mosaicism in hemophilia A: a fairly common event. *Am J Hum Genet.* 2001 Jul;69(1):75–87.
52. Oldenburg J, Rost S, El-Maarri O, Leuer M, Olek K, Müller CR, et al. De novo factor VIII gene intron 22 inversion in a female carrier presents as a somatic mosaicism. *Blood.* 2000 Oct 15;96(8):2905–6.
53. Gu W, Zhang F, Lupski JR. Mechanisms for human genomic rearrangements. *PathoGenetics.* 2008;1(1):4.
54. Stankiewicz P, Lupski JR. Genome architecture, rearrangements and genomic disorders. *Trends Genet TIG.* 2002 Feb;18(2):74–82.
55. Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet TIG.* 1998 Oct;14(10):417–22.
56. Lakich D, Kazazian HH, Antonarakis SE, Gitschier J. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nat Genet.* 1993 Nov;5(3):236–41.
57. Naylor J, Brinke A, Hassock S, Green PM, Giannelli F. Characteristic mRNA abnormality found in half the patients with severe haemophilia A is due to large DNA inversions. *Hum Mol Genet.* 1993 Nov;2(11):1773–8.
58. Antonarakis SE, Rossiter JP, Young M, Horst J, de Moerloose P, Sommer SS, et al. Factor VIII gene inversions in severe hemophilia A: results of an international consortium study. *Blood.* 1995 Sep 15;86(6):2206–12.
59. Schröder J, El-Maarri O, Schwaab R, Müller CR, Oldenburg J. Factor VIII intron-1 inversion: frequency and inhibitor prevalence. *J Thromb Haemost JTH.* 2006 May;4(5):1141–3.
60. Naylor JA, Buck D, Green P, Williamson H, Bentley D, Giannelli F. Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Hum Mol Genet.* 1995 Jul;4(7):1217–24.
61. Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, et al. The DNA sequence of the human X chromosome. *Nature.* 2005 Mar 17;434(7031):325–37.
62. Bagnall RD, Giannelli F, Green PM. Polymorphism and hemophilia A causing inversions in distal Xq28: a complex picture. *J Thromb Haemost JTH.* 2005 Nov;3(11):2598–9.
63. Bagnall RD, Giannelli F, Green PM. Int22h-related inversions causing hemophilia A: a novel insight into their origin and a new more discriminant PCR test for their detection. *J Thromb Haemost JTH.* 2006 Mar;4(3):591–8.
64. Lannoy N, Ravoet M, Grisart B, Fretigny M, Vikkula M, Hermans C. Five int22h homologous copies at the Xq28 locus identified in intron22 inversion type 3 of the Factor VIII gene. *Thromb Res.* 2016 Jan;137:224–7.

65. Carage M-L. Détection de grands réarrangements par CGH array: application à l'hémophilie A et B [Thèse d'exercice]. [Lyon, France]: Université Claude Bernard; 2010.
66. Brinke A, Tagliavacca L, Naylor J, Green P, Giangrande P, Giannelli F. Two chimaeric transcription units result from an inversion breaking intron 1 of the factor VIII gene and a region reportedly affected by reciprocal translocations in T-cell leukaemia. *Hum Mol Genet.* 1996 Dec;5(12):1945–51.
67. 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 Jan 1;99(1):168–74.
68. 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 JTH.* 2012 Aug;10(8):1600–8.
69. El-Hattab AW, Fang P, Jin W, Hughes JR, Gibson JB, Patel GS, et al. Int22h-1/int22h-2-mediated Xq28 rearrangements: intellectual disability associated with duplications and in utero male lethality with deletions. *J Med Genet.* 2011 Dec;48(12):840–50.
70. Lannoy N, Grisart B, Eeckhoudt S, Verellen-Dumoulin C, Lambert C, Vikkula M, et al. Intron 22 homologous regions are implicated in exons 1-22 duplications of the F8 gene. *Eur J Hum Genet EJHG.* 2013 Sep;21(9):970–6.
71. Li JN, Carrero IG, Dong JF, Yu FL. Complexity and diversity of F8 genetic variations in the 1000 genomes. *J Thromb Haemost JTH.* 2015 Nov;13(11):2031–40.
72. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. *Nat Rev Genet.* 2009 Aug;10(8):551–64.
73. Haviv-Chesner A, Kobayashi Y, Gabriel A, Kupiec M. Capture of linear fragments at a double-strand break in yeast. *Nucleic Acids Res.* 2007;35(15):5192–202.
74. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol.* 2017 Aug;18(8):495–506.
75. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet TIG.* 2008 Nov;24(11):529–38.
76. Roth L, Marschalek R, Oldenburg J, Oyen F, Schneppenheim R. Characterisation of two novel large F8 deletions in patients with severe haemophilia A and factor VIII inhibitors. *Thromb Haemost.* 2011 Feb;105(2):279–84.
77. Woods-Samuels P, Kazazian HH, Antonarakis SE. Nonhomologous recombination in the human genome: deletions in the human factor VIII gene. *Genomics.* 1991 May;10(1):94–101.
78. Shibata M, Shima M, Morichika S, McVey J, Tuddenham EG, Tanaka I, et al. An alloantibody recognizing the FVIII A1 domain in a patient with CRM reduced haemophilia A due to deletion of a large portion of the A1 domain DNA sequence. *Thromb Haemost.* 2000 Sep;84(3):442–8.
79. Chen C, Xie X, Wu X, Lu Y, Wang X, Wu W, et al. Complex recombination with deletion in the F8 and duplication in the TMLHE mediated by int22h copies during early embryogenesis. *Thromb Haemost.* 2017 Jul 26;117(8):1478–85.



80. You GL, Ding QL, Lu YL, Dai J, Xi XD, Wang XF, et al. Characterization of large deletions in the F8 gene using multiple competitive amplification and the genome walking technique. *J Thromb Haemost JTH*. 2013 Jun;11(6):1103–10.
81. Tavassoli K, Eigel A, Horst J. A deletion/insertion leading to the generation of a direct repeat as a result of slipped mispairing and intragenic recombination in the factor VIII gene. *Hum Genet*. 1999 May;104(5):435–7.
82. You G, Chi K, Lu Y, Ding Q, Dai J, Xi X, et al. Identification and characterisation of a novel aberrant pattern of intron 1 inversion with concomitant large insertion and deletion within the F8 gene. *Thromb Haemost*. 2014 Aug;112(2):264–70.
83. Zimmermann MA, Oldenburg J, Müller CR, Rost S. Characterization of duplication breakpoints in the factor VIII gene. *J Thromb Haemost JTH*. 2010 Dec;8(12):2696–704.
84. Aquila M, Pasino M, Lanza T, Bottini F, Molinari AC, Bicocchi MP. Duplication of exon 13 causes one third of the cases of mild hemophilia A in northern Italy. *Haematologica*. 2004 Jun;89(6):758–9.
85. Murru S, Casula L, Pecorara M, Mori P, Cao A, Pirastu M. Illegitimate recombination produced a duplication within the FVIII gene in a patient with mild hemophilia A. *Genomics*. 1990 May;7(1):115–8.
86. Aquila M, Pasino M, Lanza T, Molinari AC, Rosano C, Bicocchi MP. Exon skipping partially restores factor VIII coagulant activity in patients with mild hemophilia A with exon 13 duplication. *Haematologica*. 2005 Jul;90(7):997–9.
87. d’Oiron R, Lavergne J-M, Lavend’homme R, Benhida A, Bordet J-C, Negrier C, et al. Deletion of alanine 2201 in the FVIII C2 domain results in mild hemophilia A by impairing FVIII binding to VWF and phospholipids and destroys a major FVIII antigenic determinant involved in inhibitor development. *Blood*. 2004 Jan 1;103(1):155–7.
88. Liu M-L, Nakaya S, Thompson AR. Non-inversion factor VIII mutations in 80 hemophilia A families including 24 with alloimmune responses. *Thromb Haemost*. 2002 Feb;87(2):273–6.
89. Nakaya S, Liu ML, Thompson AR. Some factor VIII exon 14 frameshift mutations cause moderately severe haemophilia A. *Br J Haematol*. 2001 Dec;115(4):977–82.
90. Frusconi S, Passerini I, Girolami F, Masieri M, Linari S, Longo G, et al. Identification of seven novel mutations of F8C by DHPLC. *Hum Mutat*. 2002 Sep;20(3):231–2.
91. Oldenburg J, Schröder J, Schmitt C, Brackmann HH, Schwaab R. Small deletion/insertion mutations within poly-A runs of the factor VIII gene mitigate the severe haemophilia A phenotype. *Thromb Haemost*. 1998 Feb;79(2):452–3.
92. Young M, Inaba H, Hoyer LW, Higuchi M, Kazazian HH Jr, Antonarakis SE. Partial correction of a severe molecular defect in hemophilia A, because of errors during expression of the factor VIII gene. *Am J Hum Genet*. 1997 Mar;60(3):565–73.
93. Liu D, Keijzers G, Rasmussen LJ. DNA mismatch repair and its many roles in eukaryotic cells. *Mutat Res*. 2017 Jul;773:174–87.
94. Svilar D, Goellner EM, Almeida KH, Sobol RW. Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. *Antioxid Redox Signal*. 2011 Jun 15;14(12):2491–507.

95. Matsumoto N, Toga T, Hayashi R, Sugawara K, Katayanagi K, Ide H, et al. Fluorescent probes for the analysis of DNA strand scission in base excision repair. *Nucleic Acids Res.* 2010 Apr;38(7):e101.
96. Hanawalt PC, Ford JM, Lloyd DR. Functional characterization of global genomic DNA repair and its implications for cancer. *Mutat Res.* 2003 Nov;544(2-3):107-14.
97. Julsing JR, Peters GJ. Methylation of DNA repair genes and the efficacy of DNA targeted anticancer treatment. *Oncol Discov.* 2014 Oct 16;2(1):3.
98. Cooper DN, Youssoufian H. The CpG dinucleotide and human genetic disease. *Hum Genet.* 1988 Feb;78(2):151-5.
99. Lannoy N, Hermans C. Principles of genetic variations and molecular diseases: applications in hemophilia A. *Crit Rev Oncol Hematol.* 2016 Aug;104:1-8.
100. Figueiredo MS, Brownlee GG. cis-acting elements and transcription factors involved in the promoter activity of the human factor VIII gene. *J Biol Chem.* 1995 May 19;270(20):11828-38.
101. Riccardi F, Rivolta GF, Franchini M, Pattacini C, Neri TM, Tagliaferri A. Characterization of a novel mutation in the F8 promoter region associated with mild hemophilia A and resistance to DDAVP therapy. *J Thromb Haemost JTH.* 2009 Jul;7(7):1234-5.
102. Dai L, Cutler JA, Savidge GF, Mitchell MJ. Characterization of a causative mutation of hemophilia A identified in the promoter region of the factor VIII gene (F8). *J Thromb Haemost JTH.* 2008 Jan;6(1):193-5.
103. Zimmermann MA, Meier D, Oldenburg J, Müller CR, Rost S. Identification and characterization of mutations in the promoter region of the factor VIII gene. *J Thromb Haemost JTH.* 2012 Feb;10(2):314-7.
104. Nougier C, Roualdes O, Fretigny M, d'Oiron R, Costa C, Negrier C, et al. Characterization of four novel molecular changes in the promoter region of the factor VIII gene. *Haemoph Off J World Fed Hemoph.* 2014 Mar;20(2):e149-156.
105. Castaman G, Mancuso ME, Giacomelli SH, Tosetto A, Santagostino E, Mannucci PM, et al. Molecular and phenotypic determinants of the response to desmopressin in adult patients with mild hemophilia A. *J Thromb Haemost JTH.* 2009 Nov;7(11):1824-31.
106. Nance D, Fletcher SN, Bolgiano DC, Thompson AR, Josephson NC, Konkle BA. Factor VIII mutation and desmopressin-responsiveness in 62 patients with mild haemophilia A. *Haemoph Off J World Fed Hemoph.* 2013 Sep;19(5):720-6.
107. Pipe SW. Functional roles of the factor VIII B domain. *Haemoph Off J World Fed Hemoph.* 2009 Nov;15(6):1187-96.
108. Jourdy Y, Nougier C, Roualdes O, Fretigny M, Durand B, Negrier C, et al. Characterization of five associations of F8 missense mutations containing FVIII B domain mutations. *Haemoph Off J World Fed Hemoph.* 2016 Feb 25;
109. Pahl S, Pavlova A, Driesen J, Oldenburg J. Effect of F8 B domain gene variants on synthesis, secretion, activity and stability of factor VIII protein. *Thromb Haemost.* 2014 Jan;111(1):58-66.
110. Ogata K, Selvaraj SR, Miao HZ, Pipe SW. Most factor VIII B domain missense mutations are unlikely to be causative mutations for severe hemophilia A: implications for genotyping. *J Thromb Haemost JTH.* 2011 Jun;9(6):1183-90.



111. Shen BW, Spiegel PC, Chang C-H, Huh J-W, Lee J-S, Kim J, et al. The tertiary structure and domain organization of coagulation factor VIII. *Blood*. 2008 Feb 1;111(3):1240–7.
112. Jacquemin M, Lavend'homme R, Benhida A, Vanzielegem B, d'Oiron R, Lavergne JM, et al. A novel cause of mild/moderate hemophilia A: mutations scattered in the factor VIII C1 domain reduce factor VIII binding to von Willebrand factor. *Blood*. 2000 Aug 1;96(3):958–65.
113. d'Oiron R, Pipe SW, Jacquemin M. Mild/moderate haemophilia A: new insights into molecular mechanisms and inhibitor development. *Haemoph Off J World Fed Hemoph*. 2008 Jul;14 Suppl 3:138–46.
114. Liu ML, Shen BW, Nakaya S, Pratt KP, Fujikawa K, Davie EW, et al. Hemophilic factor VIII C1- and C2-domain missense mutations and their modeling to the 1.5-angstrom human C2-domain crystal structure. *Blood*. 2000 Aug 1;96(3):979–87.
115. Leyte A, van Schijndel HB, Niehrs C, Huttner WB, Verbeet MP, Mertens K, et al. Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. *J Biol Chem*. 1991 Jan 15;266(2):740–6.
116. Mumford AD, Laffan M, O'Donnell J, McVey JH, Johnson DJD, Manning RA, et al. A Tyr346->Cys substitution in the interdomain acidic region a1 of factor VIII in an individual with factor VIII:C assay discrepancy. *Br J Haematol*. 2002 Aug;118(2):589–94.
117. Pipe SW, Eickhorst AN, McKinley SH, Saenko EL, Kaufman RJ. Mild hemophilia A caused by increased rate of factor VIII A2 subunit dissociation: evidence for nonproteolytic inactivation of factor VIIIa in vivo. *Blood*. 1999 Jan 1;93(1):176–83.
118. Rudzki Z, Duncan EM, Casey GJ, Neumann M, Favalaro EJ, Lloyd JV. Mutations in a subgroup of patients with mild haemophilia A and a familial discrepancy between the one-stage and two-stage factor VIII:C methods. *Br J Haematol*. 1996 Aug;94(2):400–6.
119. Mazurier C, Gaucher C, Jorieux S, Parquet-Gernez A. Mutations in the FVIII gene in seven families with mild haemophilia A. *Br J Haematol*. 1997 Feb;96(2):426–7.
120. Pavlova A, Delev D, Pezeshkpoor B, Müller J, Oldenburg J. Haemophilia A mutations in patients with non-severe phenotype associated with a discrepancy between one-stage and chromogenic factor VIII activity assays. *Thromb Haemost*. 2014 May 5;111(5):851–61.
121. Spiegel PC, Murphy P, Stoddard BL. Surface-exposed hemophilic mutations across the factor VIII C2 domain have variable effects on stability and binding activities. *J Biol Chem*. 2004 Dec 17;279(51):53691–8.
122. Amano K, Sarkar R, Pemberton S, Kembell-Cook G, Kazazian HH, Kaufman RJ. The molecular basis for cross-reacting material-positive hemophilia A due to missense mutations within the A2-domain of factor VIII. *Blood*. 1998 Jan 15;91(2):538–48.
123. Castaman G, Giacomelli SH, Mancuso ME, D'Andrea G, Santacroce R, Sanna S, et al. Deep intronic variations may cause mild hemophilia A. *J Thromb Haemost JTH*. 2011 Aug;9(8):1541–8.
124. Pezeshkpoor B, Zimmer N, Marquardt N, Nanda I, Haaf T, Budde U, et al. Deep intronic “mutations” cause hemophilia A: application of next generation sequencing in patients without detectable mutation in F8 cDNA. *J Thromb Haemost JTH*. 2013 Sep;11(9):1679–87.

125. Bagnall RD, Waseem NH, Green PM, Colvin B, Lee C, Giannelli F. Creation of a novel donor splice site in intron 1 of the factor VIII gene leads to activation of a 191 bp cryptic exon in two haemophilia A patients. *Br J Haematol*. 1999 Dec;107(4):766–71.
126. Pezeshkpoor B, Berkemeier A-C, Czogalla KJ, Oldenburg J, El-Maarri O. Evidence of pathogenicity of a mutation in 3' untranslated region causing mild haemophilia A. *Haemoph Off J World Fed Hemoph*. 2016 Jul;22(4):598–603.
127. Kuang SQ, Hasham S, Phillips MD, Wolf D, Wan Y, Thiagarajan P, et al. Characterization of a novel autosomal dominant bleeding disorder in a large kindred from east Texas. *Blood*. 2001 Mar 15;97(6):1549–54.
128. Vincent LM, Tran S, Livaja R, Benseid TA, Milewicz DM, Dahlbäck B. Coagulation factor V(A2440G) causes east Texas bleeding disorder via TFPI $\alpha$ . *J Clin Invest*. 2013 Sep;123(9):3777–87.
129. Cunha MLR, Bakhtiari K, Peter J, Marquart JA, Meijers JCM, Middeldorp S. A novel mutation in the F5 gene (factor V Amsterdam) associated with bleeding independent of factor V procoagulant function. *Blood*. 2015 Mar 12;125(11):1822–5.
130. Langdown J, Luddington RJ, Huntington JA, Baglin TP. A hereditary bleeding disorder resulting from a premature stop codon in thrombomodulin (p.Cys537Stop). *Blood*. 2014 Sep 18;124(12):1951–6.
131. Dargaud Y, Scoazec JY, Wielders SJH, Trzeciak C, Hackeng TM, Négrier C, et al. Characterization of an autosomal dominant bleeding disorder caused by a thrombomodulin mutation. *Blood*. 2015 Feb 26;125(9):1497–501.
132. Maclachlan A, Dolan G, Grimley C, Watson SP, Morgan NV, On Behalf Of The Uk Gapp Study Group null. Whole exome sequencing identifies a mutation in thrombomodulin as the genetic cause of a suspected platelet disorder in a family with normal platelet function. *Platelets*. 2017 Sep;28(6):611–3.
133. Burley K, Whyte CS, Westbury SK, Walker M, Stirrups KE, Turro E, et al. Altered fibrinolysis in autosomal dominant thrombomodulin-associated coagulopathy. *Blood*. 2016 Oct 6;128(14):1879–83.
134. Dargaud Y, Jourdy Y, Le Quellec S, Hemker CH, Lindhout T, Castoldi E, et al. Effect of five therapeutic strategies on the coagulation defect induced by the thrombomodulin c.1611C>A mutation. *Br J Haematol*. 2016 Sep;174(6):993–6.
135. 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 May;3(5):1053–60.
136. 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 JTH*. 2013 Jul;11(7):1329–40.
137. Hsu T-C, Nakaya SM, Thompson AR. Severe haemophilia B due to a 6 kb factor IX gene deletion including exon 4: non-homologous recombination associated with a shortened transcript from whole blood. *Thromb Haemost*. 2007 Feb;97(2):176–80.
138. Chen SH, Scott CR. Recombination between two 14-bp homologous sequences as the mechanism for gene deletion in factor IX Seattle 1. *Am J Hum Genet*. 1990 Dec;47(6):1020–2.

139. Ludwig M, Grimm T, Brackmann HH, Olek K. Parental origin of factor IX gene mutations, and their distribution in the gene. *Am J Hum Genet.* 1992 Jan;50(1):164–73.
140. Ketterling RP, Vielhaber EL, Lind TJ, Thorland EC, Sommer SS. The rates and patterns of deletions in the human factor IX gene. *Am J Hum Genet.* 1994 Feb;54(2):201–13.
141. Li X, Drost JB, Roberts S, Kasper C, Sommer SS. Factor IX mutations in South Africans and African Americans are compatible with primarily endogenous influences upon recent germline mutations. *Hum Mutat.* 2000 Oct;16(4):371.
142. Green PM, Bentley DR, Mibashan RS, Giannelli F. Partial deletion by illegitimate recombination of the factor IX gene in a haemophilia B family with two inhibitor patients. *Mol Biol Med.* 1988 Apr;5(2):95–106.
143. Solera J, Magallón M, Martín-Villar J, Coloma A. Factor IX Madrid 2: a deletion/insertion in factor IX gene which abolishes the sequence of the donor junction at the exon IV-intron d splice site. *Am J Hum Genet.* 1992 Feb;50(2):434–7.
144. Ketterling RP, Ricke DO, Wurster MW, Sommer SS. Deletions with inversions: report of a mutation and review of the literature. *Hum Mutat.* 1993;2(1):53–7.
145. Wu X, Lu Y, Ding Q, You G, Dai J, Xi X, et al. Characterisation of large F9 deletions in seven unrelated patients with severe haemophilia B. *Thromb Haemost.* 2014 Sep 2;112(3):459–65.
146. Venceslá A, Barceló MJ, Baena M, Quintana M, Baiget M, Tizzano EF. Marker and real-time quantitative analyses to confirm hemophilia B carrier diagnosis of a complete deletion of the F9 gene. *Haematologica.* 2007 Nov;92(11):1583–4.
147. Alatzoglou KS, Azriyanti A, Rogers N, Ryan F, Curry N, Noakes C, et al. SOX3 deletion in mouse and human is associated with persistence of the craniopharyngeal canal. *J Clin Endocrinol Metab.* 2014 Dec;99(12):E2702-2708.
148. Hewitt J, Chou EMH, Brown LA, Smith VC, Yong SL, Wadsworth LD, et al. Molecular characterization of a 4,409,480 bp deletion of the human X chromosome in a patient with haemophilia B. *Haemoph Off J World Fed Hemoph.* 2014 May;20(3):e230-234.
149. Rousseau F, Vincent A, Rivella S, Heitz D, Triboli C, Maestrini E, et al. Four chromosomal breakpoints and four new probes mark out a 10-cM region encompassing the fragile-X locus (FRAXA). *Am J Hum Genet.* 1991 Jan;48(1):108–16.
150. Helle JR, Barøy T, Misceo D, Braaten Ø, Fannemel M, Frengen E. Hyperphagia, mild developmental delay but apparently no structural brain anomalies in a boy without SOX3 expression. *Am J Med Genet A.* 2013 May;161A(5):1137–42.
151. Buchholtz M-L, Brüning A, Mylonas I, Jückstock J. Epigenetic silencing of the LDOC1 tumor suppressor gene in ovarian cancer cells. *Arch Gynecol Obstet.* 2014 Jul;290(1):149–54.
152. Stevanović M, Lovell-Badge R, Collignon J, Goodfellow PN. SOX3 is an X-linked gene related to SRY. *Hum Mol Genet.* 1993 Dec;2(12):2013–8.
153. Rizzoti K, Brunelli S, Carmignac D, Thomas PQ, Robinson IC, Lovell-Badge R. SOX3 is required during the formation of the hypothalamo-pituitary axis. *Nat Genet.* 2004 Mar;36(3):247–55.

154. Wu Q-F, Yang L, Li S, Wang Q, Yuan X-B, Gao X, et al. Fibroblast growth factor 13 is a microtubule-stabilizing protein regulating neuronal polarization and migration. *Cell*. 2012 Jun 22;149(7):1549–64.
155. Brown SSG, Stanfield AC. Fragile X premutation carriers: A systematic review of neuroimaging findings. *J Neurol Sci*. 2015 Mar 27;
156. Plug I, Mauser-Bunschoten EP, Bröcker-Vriends AHJT, van Amstel HKP, van der Bom JG, van Diemen-Homan JEM, et al. Bleeding in carriers of hemophilia. *Blood*. 2006 Jul 1;108(1):52–6.
157. Mauser Bunschoten EP, van Houwelingen JC, Sjamsoedin Visser EJ, van Dijken PJ, Kok AJ, Sixma JJ. Bleeding symptoms in carriers of hemophilia A and B. *Thromb Haemost*. 1988 Jun 16;59(3):349–52.
158. Lyon MF. Sex chromatin and gene action in the mammalian X-chromosome. *Am J Hum Genet*. 1962 Jun;14:135–48.
159. Schröder W, Wehnert M, Herrmann FH. Intron 22 of factor VIII gene--a hot spot for structural aberrations causing severe hemophilia A. *Blood*. 1996 Apr 1;87(7):3067–8.
160. Abou-Elew H, Ahmed H, Raslan H, Abdelwahab M, Hammoud R, Mokhtar D, et al. Genotyping of intron 22-related rearrangements of F8 by inverse-shifting PCR in Egyptian hemophilia A patients. *Ann Hematol*. 2011 May;90(5):579–84.
161. Andrikovics H, Klein I, Bors A, Nemes L, Marosi A, Váradi A, et al. Analysis of large structural changes of the factor VIII gene, involving intron 1 and 22, in severe hemophilia A. *Haematologica*. 2003 Jul;88(7):778–84.
162. Arruda VR, Pieneman WC, Reitsma PH, Deutz-Terlouw PP, Annichino-Bizzacchi JM, Briët E, et al. Eleven novel mutations in the factor VIII gene from Brazilian hemophilia A patients. *Blood*. 1995 Oct 15;86(8):3015–20.
163. Strmecki L, Benedik-Dolnicar M, Vidan-Jeras B, Komel R. Inversions of the factor VIII gene in Slovenian patients with severe haemophilia A. *Eur J Haematol*. 1999 Jul;63(1):64–6.
164. Ljung R, Sjörin E. Inversions of the factor VIII gene in Swedish patients with severe haemophilia A. *Eur J Haematol*. 1995 May;54(5):310–3.
165. Naylor JA, Nicholson P, Goodeve Anne null, Hassock S, Peake I, Giannelli F. A novel DNA inversion causing severe hemophilia A. *Blood*. 1996 Apr 15;87(8):3255–61.
166. Enayat MS, Theophilus BD, Williams MD, Hill FG. Another unique variant pattern of intron 22 rearrangement in factor VIII gene seen in a hemophilia A family. *Blood*. 1995 May 1;85(9):2639.
167. Windsor S, Taylor SA, Lillicrap D. Direct detection of a common inversion mutation in the genetic diagnosis of severe hemophilia A. *Blood*. 1994 Oct 1;84(7):2202–5.
168. Mühle C, Zenker M, Chuzhanova N, Schneider H. Recurrent inversion with concomitant deletion and insertion events in the coagulation factor VIII gene suggests a new mechanism for X-chromosomal rearrangements causing hemophilia A. *Hum Mutat*. 2007 Oct;28(10):1045.
169. Fujita J, Miyawaki Y, Suzuki A, Maki A, Okuyama E, Murata M, et al. A possible mechanism for Inv22-related F8 large deletions in severe hemophilia A patients with high responding factor VIII inhibitors. *J Thromb Haemost JTH*. 2012 Oct;10(10):2099–107.

170. Isrie M, Froyen G, Devriendt K, de Ravel T, Fryns JP, Vermeesch JR, et al. Sporadic male patients with intellectual disability: contribution of X-chromosome copy number variants. *Eur J Med Genet.* 2012 Nov;55(11):577–85.
171. Vanmarsenille L, Giannandrea M, Fieremans N, Verbeeck J, Belet S, Raynaud M, et al. Increased dosage of RAB39B affects neuronal development and could explain the cognitive impairment in male patients with distal Xq28 copy number gains. *Hum Mutat.* 2014 Mar;35(3):377–83.
172. Miskinyte S, Butler MG, Hervé D, Sarret C, Nicolino M, Petralia JD, et al. Loss of BRCC3 deubiquitinating enzyme leads to abnormal angiogenesis and is associated with syndromic moyamoya. *Am J Hum Genet.* 2011 Jun 10;88(6):718–28.
173. Scott RM, Smith ER. Moyamoya disease and moyamoya syndrome. *N Engl J Med.* 2009 Mar 19;360(12):1226–37.
174. Janczar S, Fogtman A, Koblowska M, Baranska D, Pastorczak A, Wegner O, et al. Novel severe hemophilia A and moyamoya (SHAM) syndrome caused by Xq28 deletions encompassing F8 and BRCC3 genes. *Blood.* 2014 Jun 19;123(25):4002–4.
175. Janczar S, Kosinska J, Ploski R, Pastorczak A, Wegner O, Zalewska-Szewczyk B, et al. Haemophilia A and cardiovascular morbidity in a female SHAM syndrome carrier due to skewed X chromosome inactivation. *Eur J Med Genet.* 2016 Jan;59(1):43–7.
176. Oldenburg J, Pavlova A. Genetic risk factors for inhibitors to factors VIII and IX. *Haemoph Off J World Fed Hemoph.* 2006 Dec;12 Suppl 6:15–22.
177. Baralle D, Baralle M. Splicing in action: assessing disease causing sequence changes. *J Med Genet.* 2005 Oct;42(10):737–48.
178. Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol.* 2011 Jul 1;3(7).
179. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell.* 2009 Feb 20;136(4):701–18.
180. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet.* 2008 Dec;40(12):1413–5.
181. Maniatis T, Tasic B. Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature.* 2002 Jul 11;418(6894):236–43.
182. Burge CB, Karlin S. Finding the genes in genomic DNA. *Curr Opin Struct Biol.* 1998 Jun;8(3):346–54.
183. Carmel I, Tal S, Vig I, Ast G. Comparative analysis detects dependencies among the 5' splice-site positions. *RNA N Y N.* 2004 May;10(5):828–40.
184. Reddy R, Henning D, Busch H. Pseudouridine residues in the 5'-terminus of uridine-rich nuclear RNA I (U1 RNA). *Biochem Biophys Res Commun.* 1981 Feb 27;98(4):1076–83.
185. Madhani HD, Guthrie C. Dynamic RNA-RNA interactions in the spliceosome. *Annu Rev Genet.* 1994;28:1–26.



186. Wu S, Romfo CM, Nilsen TW, Green MR. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*. 1999 Dec 16;402(6763):832–5.
187. Gao K, Masuda A, Matsuura T, Ohno K. Human branch point consensus sequence is yUnAy. *Nucleic Acids Res*. 2008 Apr;36(7):2257–67.
188. Kol G, Lev-Maor G, Ast G. Human-mouse comparative analysis reveals that branch-site plasticity contributes to splicing regulation. *Hum Mol Genet*. 2005 Jun 1;14(11):1559–68.
189. Berget SM. Exon recognition in vertebrate splicing. *J Biol Chem*. 1995 Feb 10;270(6):2411–4.
190. Hertel KJ. Combinatorial control of exon recognition. *J Biol Chem*. 2008 Jan 18;283(3):1211–5.
191. Fox-Walsh KL, Dou Y, Lam BJ, Hung S-P, Baldi PF, Hertel KJ. The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proc Natl Acad Sci U S A*. 2005 Nov 8;102(45):16176–81.
192. Deshler JO, Rossi JJ. Unexpected point mutations activate cryptic 3' splice sites by perturbing a natural secondary structure within a yeast intron. *Genes Dev*. 1991 Jul;5(7):1252–63.
193. Jeong S. SR Proteins: Binders, Regulators, and Connectors of RNA. *Mol Cells*. 2017 Jan;40(1):1–9.
194. Graveley BR. Sorting out the complexity of SR protein functions. *RNA N Y N*. 2000 Sep;6(9):1197–211.
195. Wang Z, Burge CB. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA N Y N*. 2008 May;14(5):802–13.
196. Geuens T, Bouhy D, Timmerman V. The hnRNP family: insights into their role in health and disease. *Hum Genet*. 2016 Aug;135(8):851–67.
197. Ladd AN, Cooper TA. Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol*. 2002 Oct 23;3(11):reviews0008.
198. Nussinov R. Conserved signals around the 5' splice sites in eukaryotic nuclear precursor mRNAs: G-runs are frequent in the introns and C in the exons near both 5' and 3' splice sites. *J Biomol Struct Dyn*. 1989 Apr;6(5):985–1000.
199. Yeo G, Hoon S, Venkatesh B, Burge CB. Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc Natl Acad Sci U S A*. 2004 Nov 2;101(44):15700–5.
200. Hui J, Hung L-H, Heiner M, Schreiner S, Neumüller N, Reither G, et al. Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing. *EMBO J*. 2005 Jun 1;24(11):1988–98.
201. Hung L-H, Heiner M, Hui J, Schreiner S, Benes V, Bindereif A. Diverse roles of hnRNP L in mammalian mRNA processing: a combined microarray and RNAi analysis. *RNA N Y N*. 2008 Feb;14(2):284–96.
202. Brudno M, Gelfand MS, Spengler S, Zorn M, Dubchak I, Conboy JG. Computational analysis of candidate intron regulatory elements for tissue-specific alternative pre-mRNA splicing. *Nucleic Acids Res*. 2001 Jun 1;29(11):2338–48.

203. Jensen KB, Musunuru K, Lewis HA, Burley SK, Darnell RB. The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. *Proc Natl Acad Sci U S A*. 2000 May 23;97(11):5740–5.
204. Singh R, Valcárcel J, Green MR. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science*. 1995 May 26;268(5214):1173–6.
205. Zimmermann MA, Gehrig A, Oldenburg J, Müller CR, Rost S. Analysis of F8 mRNA in haemophilia A patients with silent mutations or presumptive splice site mutations. *Haemoph Off J World Fed Hemoph*. 2013 Mar;19(2):310–7.
206. Wion KL, Kelly D, Summerfield JA, Tuddenham EG, Lawn RM. Distribution of factor VIII mRNA and antigen in human liver and other tissues. *Nature*. 1985 Oct 24;317(6039):726–9.
207. Hollestelle MJ, Thinnis T, Crain K, Stiko A, Kruijt JK, van Berkel TJ, et al. Tissue distribution of factor VIII gene expression in vivo--a closer look. *Thromb Haemost*. 2001 Sep;86(3):855–61.
208. Liang Q, Xiang M, Lu Y, Ruan Y, Ding Q, Wang X, et al. Characterisation and quantification of F8 transcripts of ten putative splice site mutations. *Thromb Haemost*. 2014 Dec 11;113(3).
209. El-Maarri O, Herbiniaux U, Graw J, Schröder J, Terzic A, Watzka M, et al. Analysis of mRNA in hemophilia A patients with undetectable mutations reveals normal splicing in the factor VIII gene. *J Thromb Haemost JTH*. 2005 Feb;3(2):332–9.
210. Martorell L, Corrales I, Ramirez L, Parra R, Raya A, Barquinero J, et al. Molecular characterization of ten F8 splicing mutations in RNA isolated from patient's leucocytes: assessment of in silico prediction tools accuracy. *Haemoph Off J World Fed Hemoph*. 2015 Mar;21(2):249–57.
211. Castaman G, Giacomelli SH, Mancuso ME, Sanna S, Santagostino E, Rodeghiero F. F8 mRNA studies in haemophilia A patients with different splice site mutations. *Haemoph Off J World Fed Hemoph*. 2010 Sep 1;16(5):786–90.
212. David D, Tavares A, Lavinha J. Characterization of a splicing mutation in the factor VIII gene at the RNA level. *Hum Genet*. 1995 Jan;95(1):109–11.
213. 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. *Haemoph Off J World Fed Hemoph*. 2013 Jan;19(1):100–5.
214. Laurie AD, Smith MP. Effect of the F8 mutation c.1538-2A>T on pre-mRNA splicing. *Haemoph Off J World Fed Hemoph*. 2009 Nov;15(6):1348–50.
215. Yenchitsomanus P, Thanootarakul P, Akkarapatumwong V, Oranwiroon S, Pung-Amritt P, Veerakul G, et al. Mutation causing exon 15 skipping and partial exon 16 deletion in factor VIII transcript, and a method for direct mutation detection. *Haemoph Off J World Fed Hemoph*. 2001 May;7(3):335–8.
216. Tavassoli K, Eigel A, Pollmann H, Horst J. Mutational analysis of ectopic factor VIII transcripts from hemophilia A patients: identification of cryptic splice site, exon skipping and novel point mutations. *Hum Genet*. 1997 Oct;100(5–6):508–11.
217. Theophilus BD, Enayat MS, Williams MD, Hill FG. Site and type of mutations in the factor VIII gene in patients and carriers of haemophilia A. *Haemoph Off J World Fed Hemoph*. 2001 Jul;7(4):381–91.

218. Ganguly A, Dunbar T, Chen P, Godmilow L, Ganguly T. Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. *Hum Genet.* 2003 Sep;113(4):348–52.
219. Bicocchi MP, Pasino M, Lanza T, Bottini F, Molinari AC, Rosano C, et al. Ectopic mRNA analysis and molecular modelling substantiate severe haemophilia in a patient with a FVIII gene splice mutation. *Thromb Haemost.* 2005 Feb;93(2):391–2.
220. Gau JP, Hsu HC, Chau WK, Ho CH. A novel splicing acceptor mutation of the factor VIII gene producing skipping of exon 25. *Ann Hematol.* 2003 Mar;82(3):175–7.
221. Desmet F-O, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009 May;37(9):e67.
222. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol J Comput Mol Cell Biol.* 2004;11(2–3):377–94.
223. Jian X, Boerwinkle E, Liu X. In silico tools for splicing defect prediction: a survey from the viewpoint of end users. *Genet Med Off J Am Coll Med Genet.* 2014 Jul;16(7):497–503.
224. Rooryck C, Kyndt F, Bozon D, Roux-Buisson N, Sacher F, Probst V, et al. New Family With Catecholaminergic Polymorphic Ventricular Tachycardia Linked to the Triadin Gene. *J Cardiovasc Electrophysiol.* 2015 Oct;26(10):1146–50.
225. Plantier JL, Rodriguez MH, Enjolras N, Attali O, Négrier C. A factor VIII minigene comprising the truncated intron I of factor IX highly improves the in vitro production of factor VIII. *Thromb Haemost.* 2001 Aug;86(2):596–603.
226. Matera AG, Wang Z. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol.* 2014 Feb;15(2):108–21.
227. Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, et al. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Hum Mutat.* 2012 Aug;33(8):1228–38.
228. Moehle EA, Braberg H, Krogan NJ, Guthrie C. Adventures in time and space: splicing efficiency and RNA polymerase II elongation rate. *RNA Biol.* 2014;11(4):313–9.
229. Krawczak M, Thomas NST, Hundrieser B, Mort M, Wittig M, Hampe J, et al. Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum Mutat.* 2007 Feb;28(2):150–8.
230. Roca X, Olson AJ, Rao AR, Enerly E, Kristensen VN, Børresen-Dale A-L, et al. Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics. *Genome Res.* 2008 Jan;18(1):77–87.
231. Le Guédard-Méreuze S, Vaché C, Molinari N, Vaudaine J, Claustres M, Roux A-F, et al. Sequence contexts that determine the pathogenicity of base substitutions at position +3 of donor splice-sites. *Hum Mutat.* 2009 Sep;30(9):1329–39.
232. Roca X, Krainer AR, Eperon IC. Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev.* 2013 Jan 15;27(2):129–44.



233. Smith CW, Porro EB, Patton JG, Nadal-Ginard B. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature*. 1989 Nov 16;342(6247):243–7.
234. Kim S, Cho C-S, Han K, Lee J. Structural Variation of Alu Element and Human Disease. *Genomics Inform*. 2016 Sep;14(3):70–7.
235. Roy-Engel AM. LINEs, SINEs and other retroelements: do birds of a feather flock together? *Front Biosci Landmark Ed*. 2012 Jan 1;17:1345–61.
236. Ade C, Roy-Engel AM, Deininger PL. Alu elements: an intrinsic source of human genome instability. *Curr Opin Virol*. 2013 Dec;3(6):639–45.
237. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001 Feb 15;409(6822):860–921.
238. Weiner AM, Deininger PL, Efstratiadis A. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu Rev Biochem*. 1986;55:631–61.
239. Deininger P. Alu elements: know the SINEs. *Genome Biol*. 2011 Dec 28;12(12):236.
240. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. *Nat Rev Genet*. 2002 May;3(5):370–9.
241. Comeaux MS, Roy-Engel AM, Hedges DJ, Deininger PL. Diverse cis factors controlling Alu retrotransposition: what causes Alu elements to die? *Genome Res*. 2009 Apr;19(4):545–55.
242. Roy AM, West NC, Rao A, Adhikari P, Alemán C, Barnes AP, et al. Upstream flanking sequences and transcription of SINEs. *J Mol Biol*. 2000 Sep 8;302(1):17–25.
243. Bennett EA, Keller H, Mills RE, Schmidt S, Moran JV, Weichenrieder O, et al. Active Alu retrotransposons in the human genome. *Genome Res*. 2008 Dec;18(12):1875–83.
244. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet*. 2009 Oct;10(10):691–703.
245. Deininger PL, Batzer MA. Alu repeats and human disease. *Mol Genet Metab*. 1999 Jul;67(3):183–93.
246. Shen MR, Batzer MA, Deininger PL. Evolution of the master Alu gene(s). *J Mol Evol*. 1991 Oct;33(4):311–20.
247. Mills RE, Bennett EA, Iskow RC, Luttig CT, Tsui C, Pittard WS, et al. Recently mobilized transposons in the human and chimpanzee genomes. *Am J Hum Genet*. 2006 Apr;78(4):671–9.
248. Ayarpadikannan S, Lee H-E, Han K, Kim H-S. Transposable element-driven transcript diversification and its relevance to genetic disorders. *Gene*. 2015 Mar 10;558(2):187–94.
249. Lev-Maor G, Sorek R, Levanon EY, Paz N, Eisenberg E, Ast G. RNA-editing-mediated exon evolution. *Genome Biol*. 2007;8(2):R29.
250. Athanasiadis A, Rich A, Maas S. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol*. 2004 Dec;2(12):e391.

251. Chen C, Ara T, Gautheret D. Using Alu elements as polyadenylation sites: A case of retroposon exaptation. *Mol Biol Evol.* 2009 Feb;26(2):327–34.
252. Sen SK, Han K, Wang J, Lee J, Wang H, Callinan PA, et al. Human genomic deletions mediated by recombination between Alu elements. *Am J Hum Genet.* 2006 Jul;79(1):41–53.
253. Hancks DC, Kazazian HH. Roles for retrotransposon insertions in human disease. *Mob DNA.* 2016;7:9.
254. Sukarova E, Dimovski AJ, Tchacarova P, Petkov GH, Efremov GD. An Alu insert as the cause of a severe form of hemophilia A. *Acta Haematol.* 2001;106(3):126–9.
255. Green PM, Bagnall RD, Waseem NH, Giannelli F. Haemophilia A mutations in the UK: results of screening one-third of the population. *Br J Haematol.* 2008 Oct;143(1):115–28.
256. Vidaud D, Vidaud M, Bahnak BR, Siguret V, Gispert Sanchez S, Laurian Y, et al. Haemophilia B due to a de novo insertion of a human-specific Alu subfamily member within the coding region of the factor IX gene. *Eur J Hum Genet EJHG.* 1993;1(1):30–6.
257. Wulff K, Gazda H, Schröder W, Robicka-Milewska R, Herrmann FH. Identification of a novel large F9 gene mutation—an insertion of an Alu repeated DNA element in exon e of the factor 9 gene. *Hum Mutat.* 2000 Mar;15(3):299.
258. Vorechovsky I. Transposable elements in disease-associated cryptic exons. *Hum Genet.* 2010 Feb;127(2):135–54.
259. Bogdanova N, Markoff A, Eisert R, Wermes C, Pollmann H, Todorova A, et al. Spectrum of molecular defects and mutation detection rate in patients with mild and moderate hemophilia A. *Hum Mutat.* 2007 Jan;28(1):54–60.
260. Buratti E, Chivers M, Královicová J, Romano M, Baralle M, Krainer AR, et al. Aberrant 5' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res.* 2007;35(13):4250–63.
261. Vorechovský I. Aberrant 3' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res.* 2006;34(16):4630–41.
262. Ishii S, Nakao S, Minamikawa-Tachino R, Desnick RJ, Fan J-Q. Alternative splicing in the alpha-galactosidase A gene: increased exon inclusion results in the Fabry cardiac phenotype. *Am J Hum Genet.* 2002 Apr;70(4):994–1002.
263. De Klein A, Riegman PH, Bijlsma EK, Heldoorn A, Muijtjens M, den Bakker MA, et al. A G->A transition creates a branch point sequence and activation of a cryptic exon, resulting in the hereditary disorder neurofibromatosis 2. *Hum Mol Genet.* 1998 Mar;7(3):393–8.
264. Sorek R, Lev-Maor G, Reznik M, Dagan T, Belinky F, Graur D, et al. Minimal conditions for exonization of intronic sequences: 5' splice site formation in alu exons. *Mol Cell.* 2004 Apr 23;14(2):221–31.
265. Lev-Maor G, Sorek R, Shomron N, Ast G. The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. *Science.* 2003 May 23;300(5623):1288–91.
266. Zarnack K, König J, Tajnik M, Martincorena I, Eustermann S, Stévant I, et al. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell.* 2013 Jan 31;152(3):453–66.

267. Cieniková Z, Damberger FF, Hall J, Allain FH-T, Maris C. Structural and mechanistic insights into poly(uridine) tract recognition by the hnRNP C RNA recognition motif. *J Am Chem Soc.* 2014 Oct 15;136(41):14536–44.
268. König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, et al. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol.* 2010 Jul;17(7):909–15.
269. Kim J-W, Park S-Y, Kim Y-M, Kim J-M, Kim D-J, Ryu H-M. Identification of new dinucleotide-repeat polymorphisms in factor VIII gene using fluorescent PCR. *Haemoph Off J World Fed Hemoph.* 2005 Jan;11(1):38–42.
270. Garagiola I, Seregni S, Mortarino M, Mancuso ME, Fasulo MR, Notarangelo LD, et al. A recurrent F8 mutation (c.6046C>T) causing hemophilia A in 8% of northern Italian patients: evidence for a founder effect. *Mol Genet Genomic Med.* 2016 Mar;4(2):152–9.
271. Sela N, Mersch B, Gal-Mark N, Lev-Maor G, Hotz-Wagenblatt A, Ast G. Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome. *Genome Biol.* 2007;8(6):R127.
272. Meili D, Kralovicova J, Zagalak J, Bonafé L, Fiori L, Blau N, et al. Disease-causing mutations improving the branch site and polypyrimidine tract: pseudoexon activation of LINE-2 and antisense Alu lacking the poly(T)-tail. *Hum Mutat.* 2009 May;30(5):823–31.
273. Winter PC, Egan H, McNulty O, Jones FGC, O'Donnell J, Jenkins PV. A recurrent F8 mutation in Irish haemophilia A patients: evidence for a founder effect. *Haemoph Off J World Fed Hemoph.* 2008 Mar;14(2):394–5.
274. Xie YG, Zheng H, Leggo J, Scully MF, Lillicrap D. A founder factor VIII mutation, valine 2016 to alanine, in a population with an extraordinarily high prevalence of mild hemophilia A. *Thromb Haemost.* 2002 Jan;87(1):178–9.
275. Esmon CT, Owen WG. Identification of an endothelial cell cofactor for thrombin-catalyzed activation of protein C. *Proc Natl Acad Sci U S A.* 1981 Apr;78(4):2249–52.
276. Maruyama I, Bell CE, Majerus PW. Thrombomodulin is found on endothelium of arteries, veins, capillaries, and lymphatics, and on syncytiotrophoblast of human placenta. *J Cell Biol.* 1985 Aug;101(2):363–71.
277. Conway EM. Thrombomodulin and its role in inflammation. *Semin Immunopathol.* 2012 Jan;34(1):107–25.
278. Bourin MC, Lundgren-Akerlund E, Lindahl U. Isolation and characterization of the glycosaminoglycan component of rabbit thrombomodulin proteoglycan. *J Biol Chem.* 1990 Sep 15;265(26):15424–31.
279. Conway EM, Nowakowski B, Steiner-Mosonyi M. Thrombomodulin lacking the cytoplasmic domain efficiently internalizes thrombin via nonclathrin-coated, pit-mediated endocytosis. *J Cell Physiol.* 1994 Feb;158(2):285–98.
280. Ohlin A-K, Larsson K, Hansson M. Soluble thrombomodulin activity and soluble thrombomodulin antigen in plasma. *J Thromb Haemost JTH.* 2005 May;3(5):976–82.
281. Boehme MW, Deng Y, Raeth U, Bierhaus A, Ziegler R, Stremmel W, et al. Release of thrombomodulin from endothelial cells by concerted action of TNF-alpha and neutrophils: in vivo and in vitro studies. *Immunology.* 1996 Jan;87(1):134–40.

282. Lohi O, Urban S, Freeman M. Diverse substrate recognition mechanisms for rhomboids; thrombomodulin is cleaved by Mammalian rhomboids. *Curr Biol CB*. 2004 Feb 3;14(3):236–41.
283. Menschikowski M, Hagelgans A, Eisenhofer G, Tiebel O, Siegert G. Reducing agents induce thrombomodulin shedding in human endothelial cells. *Thromb Res*. 2010 Aug;126(2):e88-93.
284. Wang L, Jiang R, Sun X-L. Recombinant thrombomodulin of different domains for pharmaceutical, biomedical, and cell transplantation applications. *Med Res Rev*. 2014 May;34(3):479–502.
285. Ohdama S, Yoshizawa Y, Kubota T, Aoki N. Plasma thrombomodulin as an indicator of thromboembolic disease in systemic lupus erythematosus. *Int J Cardiol*. 1994 Dec;47(1 Suppl):S1-6.
286. Taylor FB Jr, Peer GT, Lockhart MS, Ferrell G, Esmon CT. Endothelial cell protein C receptor plays an important role in protein C activation in vivo. *Blood*. 2001 Mar 15;97(6):1685–8.
287. Stearns-Kurosawa DJ, Kurosawa S, Mollica JS, Ferrell GL, Esmon CT. The endothelial cell protein C receptor augments protein C activation by the thrombin-thrombomodulin complex. *Proc Natl Acad Sci U S A*. 1996 Sep 17;93(19):10212–6.
288. Slungaard A, Fernandez JA, Griffin JH, Key NS, Long JR, Piegors DJ, et al. Platelet factor 4 enhances generation of activated protein C in vitro and in vivo. *Blood*. 2003 Jul 1;102(1):146–51.
289. Tsiang M, Lentz SR, Sadler JE. Functional domains of membrane-bound human thrombomodulin. EGF-like domains four to six and the serine/threonine-rich domain are required for cofactor activity. *J Biol Chem*. 1992 Mar 25;267(9):6164–70.
290. Koyama T, Parkinson JF, Sié P, Bang NU, Müller-Berghaus G, Preissner KT. Different glycoforms of human thrombomodulin. Their glycosaminoglycan-dependent modulatory effects on thrombin inactivation by heparin cofactor II and antithrombin III. *Eur J Biochem*. 1991 Jun 15;198(3):563–70.
291. Bajzar L, Morser J, Nesheim M. TAFI, or plasma procarboxypeptidase B, couples the coagulation and fibrinolytic cascades through the thrombin-thrombomodulin complex. *J Biol Chem*. 1996 Jul 12;271(28):16603–8.
292. Hosaka Y, Takahashi Y, Ishii H. Thrombomodulin in human plasma contributes to inhibit fibrinolysis through acceleration of thrombin-dependent activation of plasma procarboxypeptidase B. *Thromb Haemost*. 1998 Feb;79(2):371–7.
293. Bajzar L, Manuel R, Nesheim ME. Purification and characterization of TAFI, a thrombin-activable fibrinolysis inhibitor. *J Biol Chem*. 1995 Jun 16;270(24):14477–84.
294. Mosnier LO. Platelet factor 4 inhibits thrombomodulin-dependent activation of thrombin-activatable fibrinolysis inhibitor (TAFI) by thrombin. *J Biol Chem*. 2011 Jan 7;286(1):502–10.
295. de Munk GA, Parkinson JF, Groeneveld E, Bang NU, Rijken DC. Role of the glycosaminoglycan component of thrombomodulin in its acceleration of the inactivation of single-chain urokinase-type plasminogen activator by thrombin. *Biochem J*. 1993 Mar 15;290 (Pt 3):655–9.
296. Schenk-Braat EA, Morser J, Rijken DC. Identification of the epidermal growth factor-like domains of thrombomodulin essential for the acceleration of thrombin-mediated inactivation of single-chain urokinase-type plasminogen activator. *Eur J Biochem*. 2001 Nov;268(21):5562–9.

297. Wilhelm S, Wilhelm O, Schmitt M, Graeff H. Inactivation of receptor-bound pro-urokinase-type plasminogen activator (pro-uPA) by thrombin and thrombin/thrombomodulin complex. *Biol Chem Hoppe Seyler*. 1994 Sep;375(9):603–8.
298. Molinari A, Giorgetti C, Lansen J, Vaghi F, Orsini G, Faioni EM, et al. Thrombomodulin is a cofactor for thrombin degradation of recombinant single-chain urokinase plasminogen activator “in vitro” and in a perfused rabbit heart model. *Thromb Haemost*. 1992 Feb 3;67(2):226–32.
299. Esmon CT. Introduction: are natural anticoagulants candidates for modulating the inflammatory response to endotoxin? *Blood*. 2000 Feb 15;95(4):1113–6.
300. Esmon CT. New mechanisms for vascular control of inflammation mediated by natural anticoagulant proteins. *J Exp Med*. 2002 Sep 2;196(5):561–4.
301. Grey ST, Tsuchida A, Hau H, Orthner CL, Salem HH, Hancock WW. Selective inhibitory effects of the anticoagulant activated protein C on the responses of human mononuclear phagocytes to LPS, IFN-gamma, or phorbol ester. *J Immunol Baltim Md 1950*. 1994 Oct 15;153(8):3664–72.
302. Grey ST, Hancock WW. A physiologic anti-inflammatory pathway based on thrombomodulin expression and generation of activated protein C by human mononuclear phagocytes. *J Immunol Baltim Md 1950*. 1996 Mar 15;156(6):2256–63.
303. Joyce DE, Gelbert L, Ciaccia A, DeHoff B, Grinnell BW. Gene expression profile of antithrombotic protein c defines new mechanisms modulating inflammation and apoptosis. *J Biol Chem*. 2001 Apr 6;276(14):11199–203.
304. Riewald M, Petrovan RJ, Donner A, Mueller BM, Ruf W. Activation of endothelial cell protease activated receptor 1 by the protein C pathway. *Science*. 2002 Jun 7;296(5574):1880–2.
305. Weiler H, Isermann BH. Thrombomodulin. *J Thromb Haemost JTH*. 2003 Jul;1(7):1515–24.
306. Conway EM, Van de Wouwer M, Pollefeyt S, Jurk K, Van Aken H, De Vriese A, et al. The lectin-like domain of thrombomodulin confers protection from neutrophil-mediated tissue damage by suppressing adhesion molecule expression via nuclear factor kappaB and mitogen-activated protein kinase pathways. *J Exp Med*. 2002 Sep 2;196(5):565–77.
307. Faust SN, Levin M, Harrison OB, Goldin RD, Lockhart MS, Kondaveeti S, et al. Dysfunction of endothelial protein C activation in severe meningococcal sepsis. *N Engl J Med*. 2001 Aug 9;345(6):408–16.
308. Glaser CB, Morser J, Clarke JH, Blasko E, McLean K, Kuhn I, et al. Oxidation of a specific methionine in thrombomodulin by activated neutrophil products blocks cofactor activity. A potential rapid mechanism for modulation of coagulation. *J Clin Invest*. 1992 Dec;90(6):2565–73.
309. Conway EM, Rosenberg RD. Tumor necrosis factor suppresses transcription of the thrombomodulin gene in endothelial cells. *Mol Cell Biol*. 1988 Dec;8(12):5588–92.
310. Boehme MWJ, Galle P, Stremmel W. Kinetics of thrombomodulin release and endothelial cell injury by neutrophil-derived proteases and oxygen radicals. *Immunology*. 2002 Nov;107(3):340–9.
311. Li J, Garnette CS, Cahn M, Claytor RB, Rohrer MJ, Dobson JG, et al. Recombinant thrombomodulin inhibits arterial smooth muscle cell proliferation induced by thrombin. *J Vasc Surg*. 2000 Oct;32(4):804–13.



312. Olivot JM, Estebanell E, Lafay M, Brohard B, Aiach M, Rendu F. Thrombomodulin prolongs thrombin-induced extracellular signal-regulated kinase phosphorylation and nuclear retention in endothelial cells. *Circ Res*. 2001 Apr 13;88(7):681–7.
313. Hosaka Y, Higuchi T, Tsumagari M, Ishii H. Inhibition of invasion and experimental metastasis of murine melanoma cells by human soluble thrombomodulin. *Cancer Lett*. 2000 Dec 20;161(2):231–40.
314. Pike GN, Cumming AM, Hay CRM, Bolton-Maggs PHB, Burthem J. Sample conditions determine the ability of thrombin generation parameters to identify bleeding phenotype in FXI deficiency. *Blood*. 2015 Jul 16;126(3):397–405.
315. Trossaert M, Lienhart A, Nougier C, Fretigny M, Sigaud M, Meunier S, et al. Diagnosis and management challenges in patients with mild haemophilia A and discrepant FVIII measurements. *Haemoph Off J World Fed Hemoph*. 2014 Jul;20(4):550–8.
316. Delvaeye M, Noris M, De Vriese A, Esmon CT, Esmon NL, Ferrell G, et al. Thrombomodulin mutations in atypical hemolytic-uremic syndrome. *N Engl J Med*. 2009 Jul 23;361(4):345–57.
317. Kunz G, Ohlin A-K, Adami A, Zöller B, Svensson P, Lane DA. Naturally occurring mutations in the thrombomodulin gene leading to impaired expression and function. *Blood*. 2002 May 15;99(10):3646–53.
318. Kunz G, Ireland HA, Stubbs PJ, Kahan M, Coulton GC, Lane DA. Identification and characterization of a thrombomodulin gene mutation coding for an elongated protein with reduced expression in a kindred with myocardial infarction. *Blood*. 2000 Jan 15;95(2):569–76.
319. Stearns DJ, Kurosawa S, Esmon CT. Microthrombomodulin. Residues 310-486 from the epidermal growth factor precursor homology domain of thrombomodulin will accelerate protein C activation. *J Biol Chem*. 1989 Feb 25;264(6):3352–6.
320. Hayashida K, Bartlett AH, Chen Y, Park PW. Molecular and cellular mechanisms of ectodomain shedding. *Anat Rec Hoboken NJ* 2007. 2010 Jun;293(6):925–37.
321. Cheng T-L, Wu Y-T, Lin H-Y, Hsu F-C, Liu S-K, Chang B-I, et al. Functions of rhomboid family protease RHBDL2 and thrombomodulin in wound healing. *J Invest Dermatol*. 2011 Dec;131(12):2486–94.
322. Kurosawa S, Stearns-Kurosawa DJ, Kinasewitz GT. Soluble thrombomodulin: a sign of bad times. *Crit Care Med*. 2008 Mar;36(3):985–7.
323. Dharmasaroja P, Dharmasaroja PA, Sobhon P. Increased plasma soluble thrombomodulin levels in cardioembolic stroke. *Clin Appl Thromb Off J Int Acad Clin Appl Thromb*. 2012 Jun;18(3):289–93.
324. Mann KG. Thrombin generation in hemorrhage control and vascular occlusion. *Circulation*. 2011 Jul 12;124(2):225–35.
325. Higuchi M, Antonarakis SE, Kasch L, Oldenburg J, Economou-Petersen E, Olek K, et al. Molecular characterization of mild-to-moderate hemophilia A: detection of the mutation in 25 of 29 patients by denaturing gradient gel electrophoresis. *Proc Natl Acad Sci U S A*. 1991 Oct 1;88(19):8307–11.
326. Higuchi M, Kazazian HH, Kasch L, Warren TC, McGinniss MJ, Phillips JA, et al. Molecular characterization of severe hemophilia A suggests that about half the mutations are not within the

- coding regions and splice junctions of the factor VIII gene. *Proc Natl Acad Sci U S A*. 1991 Aug 15;88(16):7405–9.
327. 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 Aug 15;92(4):1458–9.
328. Oldenburg J, Ivaskevicius V, Rost S, Fregin A, White K, Holinski-Feder E, et al. Evaluation of DHPLC in the analysis of hemophilia A. *J Biochem Biophys Methods*. 2001 Jan 30;47(1–2):39–51.
329. Aquila M, Pasino M, Di Duca M, Bottini F, Molinari AC, Bicocchi MP. MLPA assay in F8 gene mutation screening. *Haemoph Off J World Fed Hemoph*. 2008 May;14(3):625–7.
330. Schneppenheim R, Schröder J, Obser T, Oyen F, Schneppenheim S, Oldenburg J. The problem of novel FVIII missense mutations for haemophilia A genetic counseling. *Hämostaseologie*. 2009 May;29(2):158–60.