UNIVERSITAT DE BARCELONA

Aplicació de les noves tecnologies de seqüenciació massiva al diagnòstic molecular de la malaltia de von Willebrand. Estudi de grans cohorts i anàlisi de la correlació genotip-fenotip

Nina Borràs Agustí

Departament de Genètica, Microbiologia i Estadística
Programa de Doctorat en Genètica
Facultat de Biologia
Universitat de Barcelona

Aplicació de les noves tecnologies de seqüenciació massiva al diagnòstic molecular de la malaltia de von Willebrand. Estudi de grans cohorts i anàlisi de la correlació genotip-fenotip.

Memòria presentada per
Nina Borràs Agustí

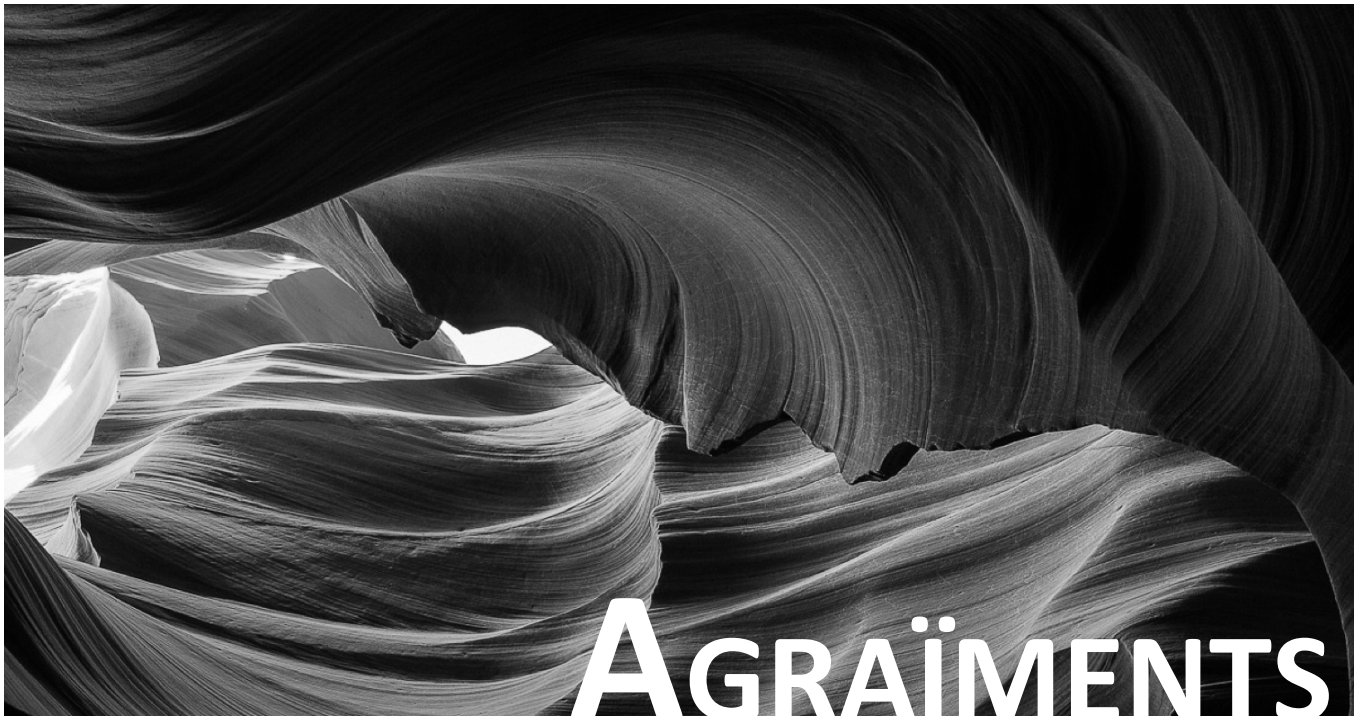
Per optar al grau de
Doctora per la Universitat de Barcelona

Tesi dirigida pel Dr. Francisco Vidal Pérez i la Dra. Irene Corrales Insa, al laboratori de Coaguolpaties Congènites del Banc de Sang i Teixits i sota la supervisió del Dr. Bru Cormand Rifà del Departament de Genètica de la Universitat de Barcelona.

Dr. Francisco Vidal Pérez Dra. Irene Corrales Insa Dr. Bru Cormand Rifà

Nina Borràs Agustí
Doctoranda

Barcelona, setembre 2018



AGRAÏMENTS

"Resulta més gratificant arribar al cim i compartir la teva experiencia amb els altres que mostrar-te tú allí sol, exhaust."

Shandel Slaten

Als meus pares, Carles i Carme

A la meva germana, Tània

AGRAÏMENTS

Si avui estic aquí és gràcies a cada una de les persones que m'han acompanyat durant aquests quatre anys. M'agradaria poder transmetre tot el que heu suposat per mi però no em surten ni les paraules, així que espero que vosaltres sapigueu més del que pugui reflectir a continuació.

En primer lloc, m'agradaria donar les gràcies als meus directors de tesi, el Dr. Francisco Vidal i la Dra. Irene Corrales per haver confiat en mi, per fer-me sentir part del grup des del primer dia i haver-me guiat en aquesta etapa. Gràcies per formar-me, per induir-me a pensar i reflexionar, i sobretot per fer-me créixer com a investigadora. Gràcies per la vostra proximitat i per no deixar-me caure en els moments de flaquesa.

Gràcies al Dr. Rafael Parra i la Dra. Carme Altisent de la Unitat d'Hemofília de la Vall d'Hebron per ser el canal d'unió amb els pacients i fer que part d'aquesta tesi hagi estat possible. També agrair a tots els investigadors del projecte PCM-EVW-ES, en especial al Dr. Javier Batlle y la Dra. Almudena Pérez. A la Dra. Teresa Fidalgo per la seva predisposició a participar i col·laborar en els estudis d'investigació de la malaltia de von Willebrand.

Gràcies a l'Associació Catalana de l'hemofília per ser la primera en apostar pel diagnòstic molecular en aquesta malaltia, per tots els pacients a qui representen, pel suport que els ofereixen i per ser un vincle entre ells i la recerca.

Gràcies al meu tutor de tesi, el Dr. Bru Cormand per la seva disponibilitat immediata per resoldre els meus dubtes, sobretot en aquesta última etapa.

Gràcies als companys del laboratori amb els que he compartit el dia a dia: treball, problemes, riures i plors (més dels que m'agradaria). A la Lorena per la seva alegria (i

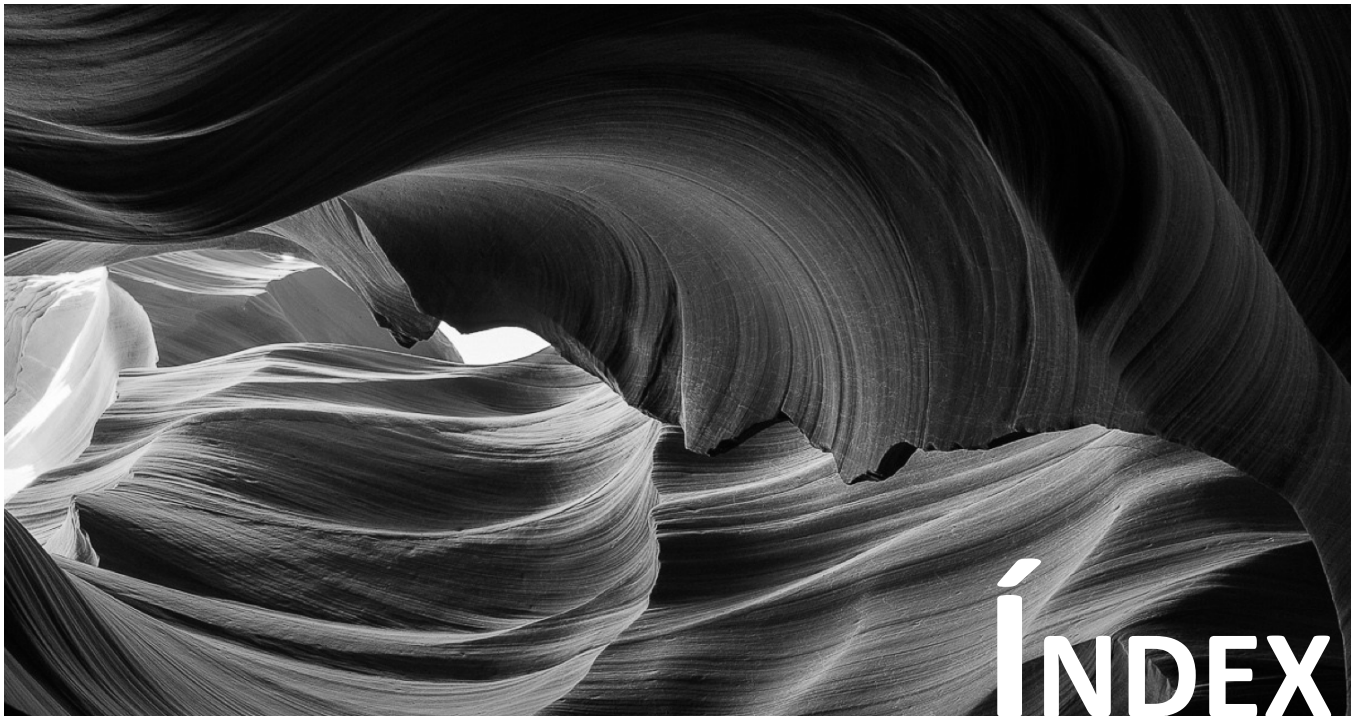
su risa inconfundible) i pels seus consells laborals i personals, gràcies per ser-hi. A la Natàlia pels *cafelitos* del migdia, molt necessaris a vegades. Al Lluís per compartir-me tot el seu saber en cultius cel·lulars. Al Gerard per ser ell mateix, per la seva calma que ha compensat en molts moments el meu nerviosisme a la sala de cultius. A elles, l'Emma, la Laura, l'Iris i la Laia perquè han estat una part indispensable durant aquest camí, gràcies per alegrar-me els dies, i per escoltar els meus drames, que no han estat pocs i animar-me tant i tan bé durant la recta final.

A les meves amigues per fer-me oblidar els problemes i riure com si no hi hagués demà, per donar-me suport, per recolzar-me i per ser-hi.

Gràcies infinitament a la meva família, als meus pares i a la Tània, pel suport incondicional en tot allò que faig, no podria sense vosaltres, sou la meva base. A tu padrina.

I, en especial, a tu, Raúl.

A tots vosaltres, gràcies.

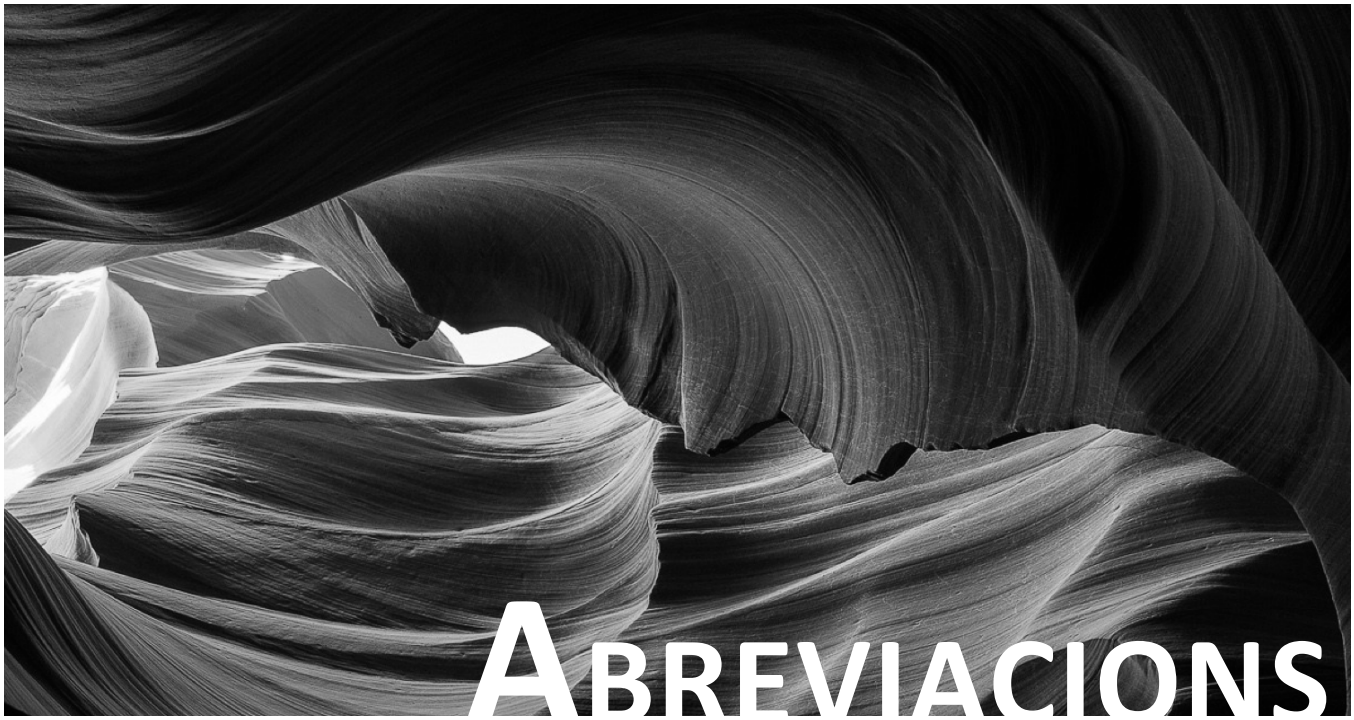


ÍNDEX

“El camí més llarg del món comença amb un pas: som-hi”

John F. Kennedy

ÍNDEX



ABREVIACIONS

“Haz las cosas lo más simple que puedas, pero no te limites a lo simple.”

Albert Einstein

ABREVIACIONS

aa: aminoàcid

ADAMTS: desintegrina i metal·loproteasa amb motius trombospondina

Ag: antigen

AINES: antiinflamatoris no esteroïdals

aPTT: temps de protrombina parcial activada

BS: puntuació de sagnat

BAT: qüestionari per a l'avaluació de sagnat

BOECs: cèl·lules endotelials derivades de sang perifèrica

CB: unió a col·lagen

cDNA: DNA complementari

DDAVP: 1-deamino-8-D-arginina vasopressina

DNA: àcid desoxiribonucleic

dNTPs: desoxinucleòtids

ddNTPs: didesoxinucleòtids

FDA: administració d'aliments i medicaments

FVIII: factor VIII de la coagulació

FIX: factor IX de la coagulació

Gp: receptor plaquetari de glicoproteïna

GWAS: estudi d'associació de genoma complet

HA: Hemofília A

HB: Hemofília B

HMWM: multímers d'alt pes molecular

HUVEC: cèl·lules endotelials de vena de cordó umbilical

indel: inserció deleció

ISTH: Societat Internacional de Trombosi i Hemostàsia

IMWM: multímers de pes molecular intermedi

ABREVIACIONS

kb: kilobase

kD: kilodalton

LMWM: multímers de baix pes molecular

LR-PCR: PCR de fragments llargs

LOVD-VWD-EAHAD: base de dades internacional de la malaltia de von Willebrand

LSMD: base de dades de mutacions específiques per a un locus

mRNA: RNA missatger

ng: nanogram

nt: nucleòtid

NGS: seqüenciació de nova generació

NMD: degradació de l'mRNA mediada per mutacions sense sentit

pb: parells de bases

PCR: reacció en cadena de la polimerasa

PSSM: potencials mutacions d'*splicing*

PTC: codó de terminació prematur

RCo: cofactor de la ristocetina

RIPA: agregació plaquetària induïda per ristocetina

RNA: àcid ribonucleic

RT-PCR: PCR en transcripció reversa

rVWF: recombinant del factor von Willebrand

SNP: polimorfisme d'un únic nucleòtid

STR: microsatèl·lit

TGS: seqüenciació de tercera generació

µg: microgram

µl: microlitre

VNTR: minisatèl·lit

VWD: malaltia de von Willebrand

VWF: factor de von Willebrand

VWF: gen del factor de von Willebrand

VWFP: pseudogèn del factor de von Willebrand

VWP: promotor del gen del factor von Willebrand

VWFpp: propèptid del factor de von Willebrand

CES: seqüenciació de l'exoma clínic

WES: seqüenciació de l'exoma complet

WGS: seqüenciació del genoma complet

Introducció

“Daria todo lo que se, por la mitad de lo que ignoro.”

Descartes

INTRODUCCIÓ

1 Història de la malaltia de von Willebrand: des de la primera descripció fins al gen responsable

La malaltia de von Willebrand (VWD, de les sigles en anglès de *Von Willebrand Disease*) es va descriure per primera vegada l'any 1926 pel doctor internista finlandès Erik Adolf von Willebrand, en una família de l'illa de Föglö, situada a l'arxipèlag Åland del Mar Bàltic.[1] El doctor von Willebrand treballava a Hèlsinki quan va visitar a una nena de 5 anys que es deia Hjördis Sundblom amb una simptomatologia d'epistaxi recurrent, hemorràgies per extraccions dentals i sagnat labial, un dels quals se li va perllongar durant 3 dies.[2] La nena formava part d'una família d'onze germans (Familia S, Sundblom (**¡Error! No se encuentra el origen de la referencia.**)), dels quals, set manifestaven

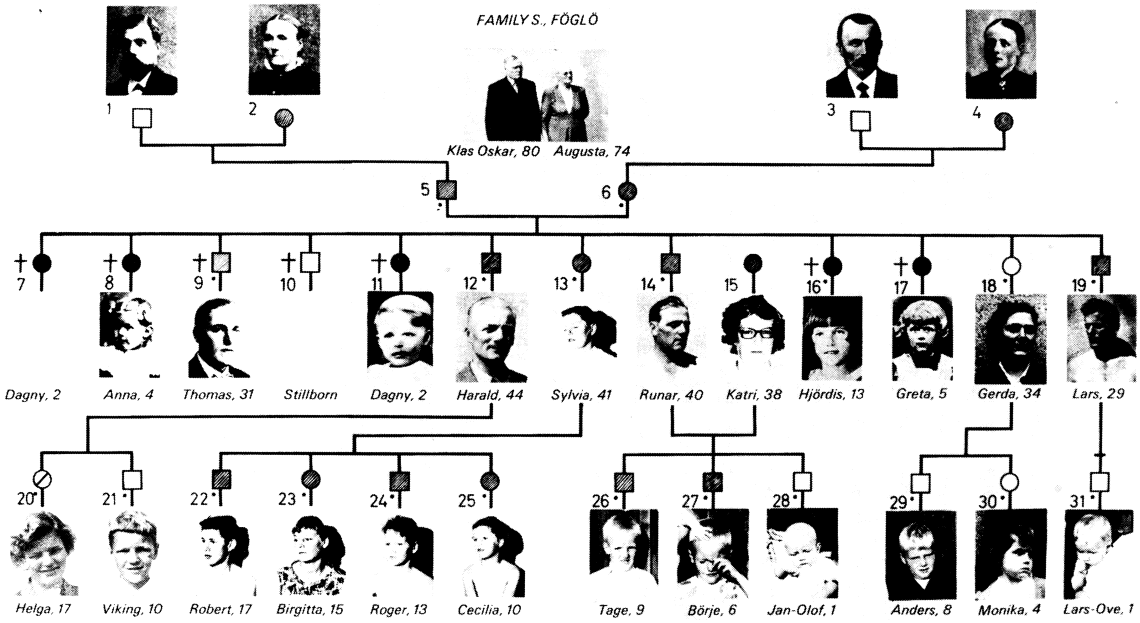


Figura 1. Arbre genealògic de la primera família diagnosticada amb la VWD. *Hjördis és la número 16 de la imatge. Quatre germanes (número 7, 8, 11 i 17) van morir degut a hemorràgies no controlades quan tenien menys de 6 anys d'edat. La resta de germans presentaven hemorràgies nasals (número 12, 13 i 14) o no presentaven simptomatologia (número 9, 10, 18 i 19).*

línica hemorràgica i, com a conseqüència dels sagnats, quatre germanes havien mort prèviament.

Els pares únicament presentaven episodis d'epistaxi. Arrel d'aquesta patologia la Hjördis va morir, durant el seu quart període menstrual, als 13 anys.

Encuriosit per la malaltia de la nena, el doctor Erik von Willebrand va ampliar l'estudi a dues famílies més (Família E i Família J), interrelacionades amb la Família S, i va descriure que 23 (16 dones i 7 homes) dels 66 individus presentaven clínica hemorràgica,[3] sent els símptomes més predominants l'epistaxi, les gingivorràgies, les menorràgies i els sagnats després d'una extracció dental o deguts a petites ferides. D'aquesta manera, va descobrir que la malaltia afectava tant a homes com a dones, mostrant un patró d'herència autosòmic dominant, diferent al descrit en pacients amb hemofília (herència lligada al cromosoma X). Addicionalment, els pacients presentaven un temps de sagnat prolongat, malgrat tenir un recompte plaquetari normal. En aquestes circumstàncies, el doctor Erik von Willebrand va considerar aquesta patologia hemorràgica una forma desconeguda d'hemofília, que va anomenar Pseudohemofília hereditària, i va postular que els sagnats podrien explicar-se per una alteració de la funció de les plaquetes i dels vasos sanguinis.[4 5]

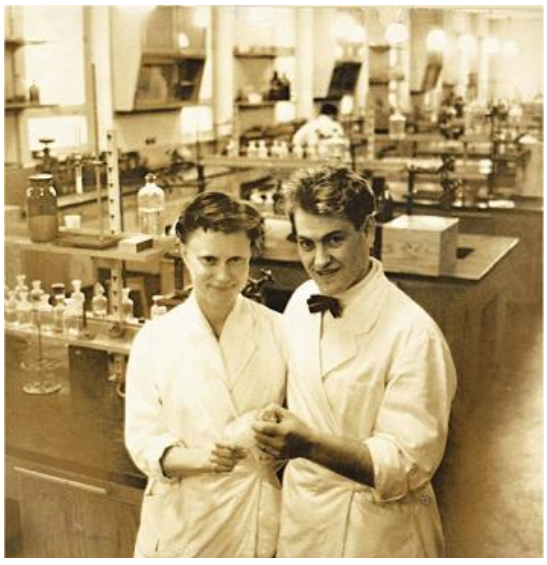


Figura 2. Margareta Blombäck i Birger Blombäck amb una ampolla de Fracció I-0 al laboratori Karolinska Institute d'Estocolm.

No va ser fins trenta anys més tard, durant la dècada dels 50, que el grup d'investigació suec format per Inga Marie Nilsson, Margareta Blombäck i Irene von Francken va demostrar que aquesta malaltia també es relacionava amb una disminució de l'activitat del FVIII de la coagulació.[6] Al mateix temps, la professora Margareta Blombäck i el seu exmarit Birger Blombäck van descobrir el primer tractament per a la malaltia, gràcies a un experiment fallit quan treballaven per obtenir un purificat de fibrinogen a partir de

la Fracció I de Cohn. El resultat d'aquest assaig va ser un precipitat de fibrinogen i FVIII que es va anomenar Fracció I-0 i va ser el tractament per als pacients amb VWD i hemofília A fins al 1980, sent disponible a Suècia una dècada abans que a la resta de països (Figura 2).[7] Posteriorment, es va observar que la infusió de concentrats de factor plasmàtic, Fracció I-0 sense FVIII o plasma preparat de pacients amb hemofília A (deficient en FVIII), podien compensar les alteracions en la coagulació dels pacients amb VWD, però no es corregia amb la infusió de fibrinogen purificat.[7] Tots aquest resultats semblaven indicar que el causant d'un temps d'hemorràgia prolongat en aquesta patologia no era un defecte funcional de les plaquetes ni dels vasos sanguinis sinó el dèficit d'un factor plasmàtic diferent al FVIII. Ara bé, no va ser fins al 1971, quan es va descobrir mitjançant proves immunològiques que es tractava d'una proteïna plasmàtica nova, que es va anomenar factor von Willebrand (VWF) en honor al descobridor de la malaltia.[8] La proteïna VWF va ser caracteritzada a nivell funcional, immunològic i bioquímic al 1975.[9] Tots aquests esdeveniments van contribuir a la implementació d'una nova prova de laboratori que permetia diagnosticar pacients amb VWD mitjançant l'antibiòtic ristocetina, obtingut de l'actinomicet *Nocardia lurida*. L'ús d'aquest antibiòtic, a concentracions de 1-1,5 mg/mL, provocava l'agregació de les plaquetes al plasma normal, però no en pacients amb VWD, on l'agregació era absent o molt disminuïda en la majoria de casos.[10]

Finalment, l'any 1985 es va clonar i caracteritzar el gen codificant pel VWF (*VWF*) per quatre laboratoris independents.[11-14] Aquest fet va permetre que el grup d'investigació suec identifiques el defecte genètic en 24 pacients amb VWD que acudien al Centre d'Hemofília d'Estocolm en els exons 18, 28, 32, 43 i 45. Anys més tard, la realització del cribratge de mutacions en les regions "*hot spot*", dels exons 18, 28, 32, 43 i 45, del gen en alguns membres vius de la família de la Hjördis Sundblom va permetre identificar el defecte genètic del primer cas descrit de VWD: la deleció d'una citosina a l'exó 18 (c.2435delC), que es tradueix en una alteració en la pauta de lectura i l'aparició d'un codó stop de forma prematura.[15, 16]

2 La malaltia de von Willebrand

La VWD és la patologia hemorràgica hereditària més freqüent a la població i està causada per dèficits quantitatius o qualitatius del VWF. Es distingeixen tres grans categories: la VWD de tipus 1 caracteritzada per una reducció parcial del VWF, la VWD de tipus 3 per una absència total de factor i la VWD de tipus 2 per defectes funcionals en aquesta proteïna.[17]

2.1 Diagnòstic clínic i fenotípic de la VWD

El diagnòstic de la VWD es basa en les manifestacions clíniques de sagnat, sobretot de tipus mucocutani, la història familiar positiva de sagnats i la detecció d'anomalies quantitatives i/o qualitatives del VWF mitjançant un conjunt extens de proves de coagulació.

2.1.1 Manifestacions clíniques

Els símptomes observats en pacients diagnosticats de VWD són molt variables depenent del tipus de VWD, dels nivells de VWF funcional, de l'edat i del sexe. En general, els més comuns són els sagnats mucocutanis i els que sorgeixen al comprometre el sistema hemostàtic, com els conseqüents de cirurgies o extraccions dentals, tot i que també s'han descrit hemorràgies al sistema nerviós central o al tracte gastrointestinal en pacients amb VWD de tipus 3, en alguns de tipus 2 i rarament en pacients de tipus 1.[18] Per contra, el sagnat articular característic dels pacients amb hemofília (hemartrosi) només s'observa en alguns casos de VWD de tipus 2N o 3 que presenten nivells molt baixos de FVIII.[19] En les dones, la menorràgia i el sagnat postpart són els més freqüents.[20] Els símptomes es poden veure modificats degut a la coexistència amb altres comorbiditats o a medicacions. Per exemple, l'aspirina o els antiinflamatoris no esteroïdals (AINEs) augmenten el risc de sagnats, i els anticonceptius orals el disminueixen.

La possibilitat de mesurar de forma objectiva i quantitativa la història hemorràgica en cada pacient és una eina molt valuosa pel diagnòstic de la VWD. Per aquest motiu, s'han desenvolupat qüestionaris estandarditzats i específics (BATs, de les sigles en anglès de *Bleeding Assesment Tools*) que permeten l'assignació d'una puntuació global (BS, de l'anglès *Bleeding Score*) corresponent amb la simptomatologia de sagnat dels pacients. Al 2005 es va publicar el primer qüestionari, anomenat *Vicenza Bleeding Questionnaire*, pel diagnòstic de la VWD de tipus 1 en adults per un grup d'investigadors de l'Hospital San Bortolo de Vicenza.[21] Aquesta primera aproximació va ser posteriorment modificada, en el marc de l'estudi *Molecular and Clinical Markers for the Diagnosis and Management of Type 1 von Willebrand Disease* (*MCMDM-1VWD),* on es va millorar la sensibilitat del qüestionari i es van establir valors de BS negatius per aquells casos que havien compromès al sistema hemostàtic i no havien presentat sagnats.[22] Posteriorment, aquest qüestionari es va simplificar en el *Condensed MCMDM-1VWD*.[23] Al 2009, es va presentar el *Pediatric Bleeding Questionnaire* on es tenia en compte l'avaluació de la simptomatologia hemorràgica en pediatria i el sagnat umbilical.[24] Finalment, el qüestionari més utilitzat actualment és el presentat per la Societat Internacional de Trombosi i Hemostàsia (ISTH), l'ISTH-BAT, que permet avaluar tant la severitat com la freqüència dels sagnats en adults i infants.[25] Es considera un BS significatiu quan és ≥4 en homes, ≥6 en dones i ≥3 en nens.[26]

2.1.2 Proves de laboratori

Les proves de laboratori utilitzades pel cribratge dels defectes hemostàtics, com el temps de tromboplastina parcial activada (aPTT), el temps de protromina (PT) o l'assaig de la funció plaquetària (PFA-100), no són suficients pel diagnòstic de la VWD.[27] Consegüentment, s'han desenvolupat tests diagnòstics dirigits a la VWD i en l'actualitat es distingeixen els tests específics i els discriminatoris. Els tests específics permeten confirmar el diagnòstic i discernir els tres grans tipus de la VWD:

- **VWF:Ag**: mesura la quantitat de proteïna (antigen) de VWF al plasma mitjançant els mètodes d'ELISA (*Enzyme-Linked Immunosorbant Assay*) o de LIA (*Latex Immunoassay*).[28, 29] El rang de normalitat és de 50-200 IU/dL.

- **VWF:RCo**: mesura la funcionalitat del VWF per aglutinar plaquetes en presència de l'antibiòtic ristocetina (1 mg/mL). El rang de normalitat és de 50-200 IU/dL. Aquest test és essencial per a la detecció dels defectes qualitatius de VWD.

- **FVIII:C**: determina l'activitat procoagulant del FVIII en el plasma. El rang de normalitat és de 50-150 IU/dL. Aquest test és important pel diagnòstic dels pacient amb VWD de tipus 2N.

En un segon nivell, els tests discriminatoris, que solen estar disponibles exclusivament en centres de referència, permeten distingir entre els diferents subtipus qualitatius 2A, 2B, 2M i 2N de la VWD. Aquests tests engloben:

- **Anàlisi del patró multimèric**: és un assaig qualitatiu que determina la distribució de la mida dels multímers del VWF al plasma mitjançant electroforesis en gels d'agarosa dodecil sulfat sòdic (SDS). En condicions normals, s'observen des de dímers fins a multímers de més de 40 unitats de VWF unides, amb mides des de 500 a 20.000 kDa (Figura 3).[30, 31] L'activitat del VWF és depenent del patró

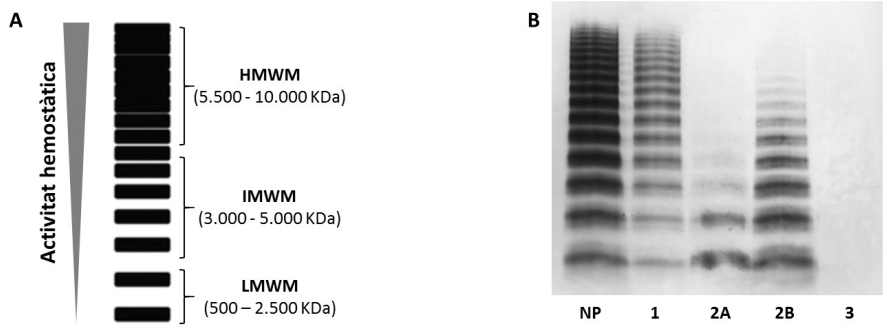


Figura 3. Representació esquemàtica de la distribució dels multímers del VWF al plasma d'un individu control i imatge dels patrons de multímers anòmals observats en els diferents tipus de VWD. *A. Els multímers d'alt pes molecular (HMWM) estan compostos per 11-20 dímers, els multímers amb pes molecular intermedi (IMWM) per 6-10 dímers i els multímers de baix pes molecular (LMWM) per 1-5 dímers. Imatge adaptada de Stockschlaeder M. et al.[31] B. Es mostren els patrons de plasma normal (NP) i dels pacients amb dèficits quantitatius (tipus 1 i 3) i qualitatius (tipus 2A i 2B) de VWD. Imatge de Sadler et al.[32]*

multimèric, sent els multímers d'alt pes molecular (HMWM, *High Molecular Weight Multimers*) els que presenten més activitat hemostàtica. Aquesta tècnica s'utilitza per diagnosticar als pacients amb VWD de tipus 2A i 2B que presenten una disminució dels HMWM.[32]

- **RIPA** (*Ristocetin Induced Platelet Aggregation*): mesura l'afinitat del VWF per agregar plaquetes a diferents concentracions de ristocetina. En concret, concentracions creixents de ristocetina (entre 0,2-1,5 mg/mL) s'afegeixen al plasma ric en plaquetes del pacient i es determina el llindar de concentració on s'indueix l'agregació plaquetària. Aquest test s'utilitza per a la detecció de pacients amb VWD de tipus 2B que agreguen a concentracions de 0,4-0,5 mg/mL, en canvi, individus control no responen a concentracions de ristocetina inferiors a 0,6-0,8 mg/mL.

- **VWF:CB**: mesura la capacitat del VWF per unir-se al col·lagen del subendoteli. Actualment únicament estan inclosos de manera rutinària en el diagnòstic de la VWD els test que mesuren l'afinitat del VWF als col·làgens de tipus I/III. Altrament, la unió a col·lagen reflexa la presència de HMWM ja que són dues mesures directament proporcionals. En conseqüència, una reducció de la unió al col·lagen reflexa la pèrdua de HMWM, o bé, una deficiència específica d'unió a col·lagen, característica específica dels pacients de tipus 2M.[30, 33] El rang de normalitat és de 50-200 IU/dL.

- **VWF:FVIIIB**: mesura l'afinitat del VWF al FVIII recombinant *in vitro* mitjançant un assaig d'ELISA. Aquest test és essencial per a la diferenciació de la VWD de tipus 2N *versus* l'hemofília A lleu.

- **VWF:pp**: mesura la quantitat de propèptid de VWF al plasma. El propèptid s'escindeix del VWF madur durant el processament intracel·lular i s'alliberen conjuntament al torrent sanguin (veure secció 3.2). En condicions normals el propèptid i el VWF es troben en una proporció 1:1 al plasma. En individus que presenten un aclariment augmentat del VWF la ràtio VWF:pp/VWF:Ag és

superior a 1, donat que aquest mecanisme tan sols afecta al VWF madur i no al propèptid.

Degut a les limitacions de les proves de laboratori i a la variabilitat dels nivells del VWF en el temps és important repetir les proves per confirmar o descartar VWD.[18, 34] La variabilitat dels nivells de VWF al plasma està influenciat per:

- Factors extrínsecs, com les situacions d'estrès, l'edat, l'exercici, els processos inflamatoris i l'embaràs. Com a exemple il·lustratiu, durant l'embaràs, l'augment d'estròg ens deriva en un augment dels nivells de VWF durant el segon i tercer trimestre ja que la síntesi d'aquest factor es troba regulada, en part, hormonalment.[35, 36]
- Factors genètics, com per exemple el grup sanguini. Individus amb grup O presenten un 25% menys de nivells de VWF que els individus de grup A, B o AB.[37] L'explicació més acceptada és la implicació del locus ABO en l'aclariment del VWF del plasma, resultant en una eliminació més ràpida del VWF i la conseg üent reducció de la seva vida mitja.[38, 39]

2.2 Classificació de la VWD

La classificació de la VWD més acceptada actualment és la del Subcomitè Científic del VWF de la ISTH, publicada per primera vegada al 1994 i revisada al 2006.[17, 32] Segons aquesta classificació, en la VWD es distingeixen dues grans categories, els defectes quantitatius que engloben la VWD de tipus 1 i 3; i els qualitatius, de tipus 2 que alhora es subclassifiquen en 2A, 2B, 2M i 2N. En la Figura 4 es mostra l'algoritme diagnòstic més emprat per a la classificació dels pacients en base als resultats obtinguts en les proves de laboratori.

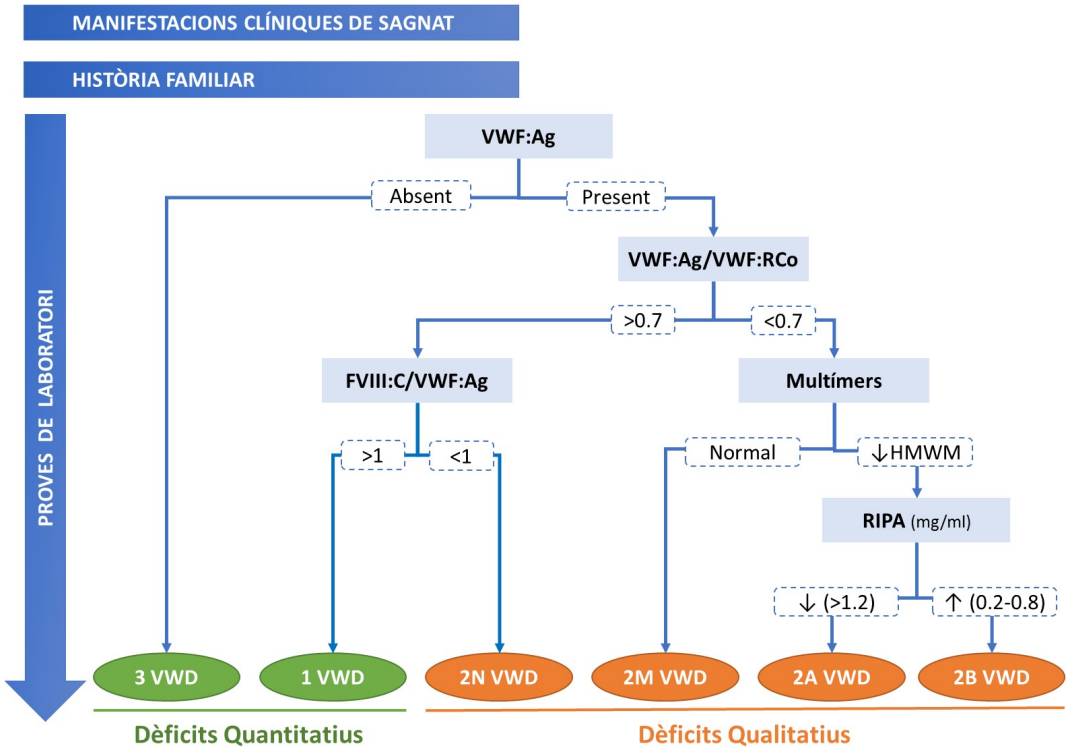


Figura 4. Representació esquemàtica de l'algoritme de diagnòstic per a la classificació dels diferents tipus de VWD. *HMWM indica multímers d'alt pes molecular.*

Per altra banda, s'ha descrit una malaltia adquirida, no congènita, que mimetitza a la VWD, i es coneguda com a síndrome de VWD adquirida (AVWS, *Acquired Von Willebrand Syndrome*).

2.2.1 Defectes quantitatius

- **VWD 1:** és el tipus més freqüent amb un 60-80% del total de casos i presenta una herència autosòmica dominant. Es caracteritza per una deficiència quantitativa parcial deguda a una reducció moderada o severa dels nivells funcionals de VWF. Presenten una relació VWF:RCo/VWF:Ag >0.7 i un patró multimèric normal, malgrat que en determinats casos es pot veure amb menys intensitat degut a la reducció de la quantitat de factor (Figura 3). Cal destacar que, els pacients amb nivells de VWF més reduïts, entre 5-30 IU/dL, tenen una simptomatologia de sagnat manifesta, mutacions dominants al *VWF* i acostuma a ser altament heretable. Per contra, pacients amb uns nivells just per sota del rang normal, 30-50 IU/dL, mostren baixa heretabilitat en el fenotip, rarament

segrega amb els símptomes de sagnat i no s'acostuma a vincula amb el *VWF*. En aquest sentit, les guies de diagnòstic de la VWD suggereixen que els pacients amb nivells >30 IU/dL haurien de ser classificats com a *Low VWF,* indicant que presenten un factor de risc de sagnat en lloc de rebre un diagnòstic específic de VWD.[17, 18] Per altra banda, dins d'aquest tipus, es distingeix una variant anomenada VWD 1C (d'aclariment) caracteritzada per un augment significatiu de l'aclariment del VWF, provocant una disminució dels nivells de factor degut a la reducció de la seva vida mitja. El diagnòstic d'aquest subtipus resulta molt important a nivell terapèutic perquè aquests pacients no són bons candidats pel tractament amb desmopressina. Aquest pacients presenten una relació VWF:pp/VWF:Ag>1.[40, 41]

- **VWD 3:** presenta una herència autosòmica recessiva (Figura 5). Es caracteritza per una absència completa de la proteïna del VWF i, com a conseqüència, uns nivells del FVIII d'entre l'1 i el 10%, ja que el VWF actua estabilitzant el FVIII en la circulació i evitant el seu aclariment (veure secció 3.3.2). Els resultats de les proves de laboratori mostren una reducció dràstica dels nivells VWF:Ag, VWF:RCo i FVIII:C. Aquests pacients, donat que no han estat mai en contacte amb el VWF, poden desenvolupar la formació d'al·loanticossos contra la proteïna del VWF durant la teràpia de substitució de factor mitjançant concentrats de VWF/FVIII (veure secció 2.4.1). La prevalença en el desenvolupament d'anticossos s'ha estimat entre 5 i el 10%.[42-44]

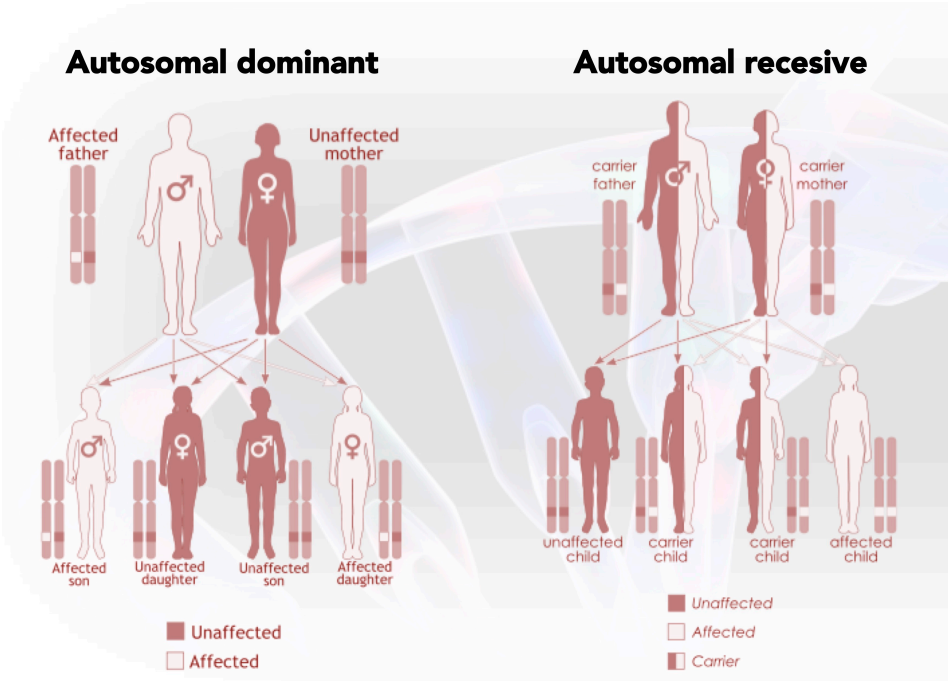


Figura 5. Representació esquemàtica del patró d'herència autosòmica dominant i recessiva.

2.2.2 Defectes qualitatius

La VWD de tipus 2 es divideix en 4 subtipus (2A, 2B, 2M, 2N) segons l'alteració funcional del VWF que afecta a la unió amb els seus lligands. En general, el patró d'herència és autosòmic dominant, excepte la VWD de tipus 2N i alguns casos de tipus 2A que són recessius. Tots els pacients de tipus 2, excepte els 2N, presenten nivells de VWF:RCo disminuïts en comparació amb els nivells de VWF:Ag, el que condueix a una relació VWF:RCo/VWF:Ag <0,7.[41]

- **VWD 2A:** és el subtipus més freqüent dels dèficits qualitatius amb un 10-15% del total de casos de VWD. Es caracteritza per una disminució de la capacitat del VWF per unir-se al receptor de les plaquetes GPIbα, degut a la pèrdua de HMWM (Figura 3). Antigament i en base al patró multimèric observat, el tipus 2A es subdividia en 4 grups: IIA, IIC, IID i IIE.[45]

- **VWD 2B:** es caracteritza per un guany de funció del VWF en la unió al receptor de les plaquetes GPIbα, fet que es posa de manifest amb l'augment del RIPA en aquests pacients. Els complexos que es formen entre els HMWM del VWF i les

plaquetes són eliminats de la circulació provocant una reducció dels HMWM lliures en plasma (Figura 3) i, en determinats casos, donant lloc a trombocitopènia.

- **VWD 2M:** es caracteritza per una reducció de la interacció del VWF amb el receptor de les plaquetes GPIbα i els lligands del teixit connectiu. Però, a diferència del subtipus 2A, no afecta al patró multimèric, característica que ha donat lloc al seu nom 2M (M, presència de HMWM). En aquest sentit, s'ha suggerit que la unió disminuïda a les plaquetes redueix l'exposició de les subunitats del VWF on es dona lloc la proteòlisi mediada per l'ADAMTS13 (*A Desintegrin And Metalloproteinasa with a ThromboSpondin type 1 motif, membre 13*), preservant així l'estructura multimèrica.[17]

- **VWD 2N**: conegut també amb el nom de Normandia, ja que és el lloc on es va descriure el primer cas.[46] Es caracteritza per una disminució de l'afinitat del VWF pel FVIII de la coagulació, donant com a resultat una disminució dels nivells de FVIII:C i una relació FVIII:C/VWF:Ag <1. Per aquest motiu, sovint es pot confondre amb l'hemofília A lleu.[47] Ambdues malalties només es poden distingir si es realitza la prova de laboratori VWF:FVIIIB.

2.2.3 Síndrome adquirida de VWD

La AVWS és una coagulopatia que mimetitza tant la simptomatologia clínica com els resultats de les proves de laboratori de la VWD congènita. No obstant, es diagnostica en pacients sense una història de sagnat ni personal ni familiar prèvia i es caracteritza per aparèixer en l'edat adulta i associada a altres malalties subjacents. Les més freqüents són els trastorns limfoproliferatius, mieloproliferatius i cardiovasculars. Els principals mecanismes patogènics de la AVWS que provoquen la disminució dels nivells plasmàtics del VWF són: la producció d'anticossos anti-VWF que augmenten l'aclariment del factor o bé n'inhibeixen la seva funció; l'absorció del VWF per cèl·lules malignes que condueix a una disminució dels nivells del VWF a la circulació; o la proteòlisi dels multímers de VWF induïda per determinats fàrmacs.[48-50]

2.2.4 Altres subtipus no contemplades per la ISTH

- **VWD 1H:** s'inclouen pacients que mostren una disminució molt lleu o, fins i tot, nivells normals de VWF però que històricament presentaven nivells compatibles amb una classificació de tipus 1. No obstant, els pacients continuen presentant episodis de sagnat.[51] Alguns autors han considerat apropiada aquesta subclassificació ja que els nivells del VWF augmenten amb l'edat, podent arribar a nivells normals.[52]

- **VWD 2A/2M:** es caracteritza per presentar nivells baixos de VWF:Ag, una disminució de la relació VWF:RCo/VWF:Ag (<0,7) i un patró multimèric característic (anomenat *smeary*), amb presència de totes les mides de multímers del VWF però amb una proporció reduïda dels HMWM.[53] Donat que aquests pacients presenten un solapament del fenotip de laboratori amb els 2A i 2M, alguns autors han proposat aquesta classificació.[62] Exemple d'aquest subtipus, són els pacients que presenten la mutació p.Arg1374Cys i han estat diagnosticats com a VWD de tipus 2A i alhora com a tipus 2M per diferents grups d'investigació.[54-59]

2.3 Epidemiologia de la VWD

La VWD engloba des de casos lleus a casos greus de simptomatologia hemorràgica. Per aquest motiu, la estimació de la seva prevalença varia molt segons el mètode de diagnòstic emprat per a la identificació dels pacients. En base als casos amb símptomes de sagnat referits a centres especialitzats s'estima que és del 0,01%.[34] No obstant, alguns pacients simptomàtics no són diagnosticats a nivell cínic i, per tant, la freqüència real de la VWD podria ser superior. En aquest sentit, la prevalença estimada a partir d'estudis epidemiològics de poblacions és d'aproximadament un 1%. En concret, al 1987, Rodeghiero *et al.* van estudiar a nivell funcional i quantitatiu el VWF en 1.218 nens (11-14 anys) procedents de varies escoles de la regió nord d'Itàlia, i es va diagnosticar a 10 nens amb VWD, donant com a resultat una prevalença de 8.200 casos per milió d'habitants (0.82%).[60] En un altre estudi, d'un total de 315 individus de

diferents grups ètnics, 8 complien criteris de VWD. Aquest estudi va concloure que la malaltia presentava una freqüència poblacional 25.000 casos per milió (2,5%) i que es troba distribuïda de forma similar a nivell mundial ja que no es van identificar diferències entre grups ètnics.[61]

En general, la distribució de la VWD per tipus s'estima que és entre 60-80% de tipus 1, 25-30% de tipus 2 i <5% de tipus 3 VWD.[17, 62] No obstant, en algunes parts d'Àsia i d'Àfrica, on el matrimoni entre membres d'una mateixa família és una pràctica habitual, la freqüència de tipus 3 VWD és més elevada.[63, 64]

L'any 2011 es va iniciar a Espanya un estudi epidemiològic d'aquesta patologia a través d'una enquesta enviada a 54 centres, entre els quals es troben les 5 unitats d'hemofília més importants. Arrel d'aquest estudi s'estima que el número de pacients amb un diagnòstic definitiu de VWD, és a dir, amb uns nivells de VWF:Ag <30 IU/dL, és de 122 casos per milió a Espanya. No obstant, si es tenen en compte aquests pacients juntament amb els que presenten un diagnòstic possible de VWD, és a dir, amb uns nivells de VWF:Ag <50 IU/dL, la prevalença incrementa fins a 207 casos per milió.[65] La distribució per tipus de VWD estimada a Espanya es mostra en la Figura 6.

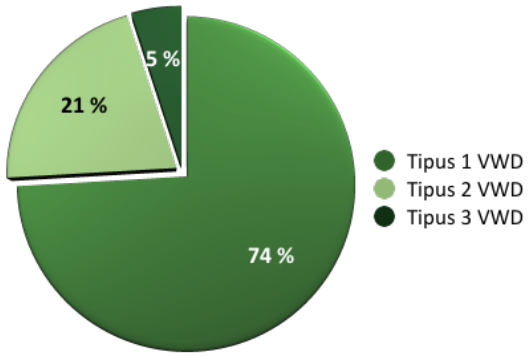


Figura 6. Estimació de la prevalença per tipus de VWD a Espanya. *La classificació s'ha establert en base als pacients que presenten un diagnòstic definitiu de VWD amb uns nivells de VWF:Ag <30%.*

Finalment, és destacable que donat que la VWD és una malaltia d'herència autosòmica s'esperaria que afectés igual a homes i dones. Malgrat això, la freqüència de dones diagnosticades amb VWD simptomàtica és sensiblement major que la d'homes, degut fonamentalment als reptes hemostàtics de la menstruació i el part.[66]

2.4 Tractament de la VWD

El tractament de la VWD es basa en controlar els sagnats i pervenir-los abans de realitzar procediments invasius com cirurgies o intervencions dentals. El tractament administrat depèn del subtipus de la VWD, la naturalesa dels sagnats i la presència d'inhibidor.[67] És destacable que la profilaxi en la VWD, a diferència del cas de l'hemofília, no és considerada com el tractament d'elecció per a la majoria de casos.[68]

2.4.1 Tractaments substitutius de VWF

- **La desmopressina** (1-deamino-8-D-arginina vasopressina o DDAVP): és un anàleg sintètic de l'hormona antidiürètica vasopressina i es va començar a utilitzar com a tractament al 1977.[69] La DDAVP actua potenciant l'alliberació endògena del VWF autòleg a través de l'estimulació de les cèl·lules endotelials, augmentant així la concentració del VWF al plasma.[70] Es pot administrar per via subcutània, nasal o intravenosa. Les concentracions de VWF/FVIII arriben al seu pic màxim al cap de 0,5-1h, produint un increment de 3 a més de 5 vegades la concentració basal de l'individu.[18] Aquest tractament és efectiu en la VWD de tipus 1, excepte en aquells casos que presenten VWF:Ag <10 UI/dL, i en la de tipus 2A, 2M i 2N VWD. [18, 71-73] En canvi, no es recomana el seu ús pel tractament de la VWD de tipus 2B atès que pot desencadenar trombocitopènia.[18, 74] Tanmateix, està contraindicat per la VWD de tipus 3, donat que els pacients no sintetitzen VWF, i en nens menors de 2 anys, degut a que presenten una baixa resposta al tractament i a que en ocasions s'ha relacionat amb hiponatrèmia en infants.[18, 75, 76]

- **Els concentrats de VWF/FVIII:** són derivats de plasma i es considera un tractament de substitució al·logènic del VWF. S'administra de forma intravenosa i és el tractament d'elecció en aquells casos on la DDAVP no és eficaç o està contraindicada.[77] Actualment, es disposa de varis concentrats de VWF/FVIII comercialitzats per diferents fabricants (Taula 1). Els primers en tenir llicència per a ser utilitzats pel tractament de pacients amb VWD van ser l'Alphanate

(Grifols) i l'Haemate-P (CSL Behring), sent aquest darrer el més utilitzat actualment.[78] Tanmateix se'n destaquen altres com el Wilfactin (LFB) que és un concentrat d'alta puresa de VWF amb una presència molt reduïda de FVIII.[79, 80]

Taula 1. Concentrats de VWF/FVIII pel tractament de la VWD.

Concentrats VWF/FVIII	Casa comercial	Activitat específica VWF*	VWF:RCo/ FVIII:C
Alphanate	Grifols, USA	>100	1,2
Biostate	SL Berhring Melbourne, Australia	100	2,0
Factor 8Y	BioProducts Laboratory, UK	>50	1,8
Fanhdi	Grifols, UK	>100	1,2
Haempate-P	CSL Berhring, USA	>80	2,4
Immunate	Baxter AG, Austria	70 ± 30	0,75
Wilate	Octapharma, USA	>100	1,0
Wilfactin	LFB, France	>80	60

**L'activitat específica de VWF de cada concentrat es mesura en Unitats Internacionals per mg (U/mg) proteïna. La ratio VWF:RCo/FVIII:C indica la concentració del VWF versus FVIII.* [79, 80]

- **Els concentrats de plaquetes:** estan indicats en aquells pacients que malgrat assoleixin nivells de VWF i FVIII normals, continuen sagnant.[68] Es consideren hemostàticament efectius ja que transporten el VWF emmagatzemat a les plaquetes (entre 10-15% del VWF total)[81] al lloc de la lesió vascular on és alliberat.[82]

2.4.2 Tractaments coadjuvants

Els tractaments coadjuvants s'utilitzen en cirurgies menors, intervencions dentals, sagnats lleus en mucoses (nasofaringi, oral o gastrointestinal) i pel tractament de la menorràgia. Es poden prescriure sols, tot i que habitualment s'administren com a complement en les teràpies de substitució amb DDAVP i concentrats de plasma. Aquestes tractaments inclouen:

- **Antifibrinolítics:** són derivats sintètics anàlegs de la lisina i formen part d'aquest grup l'àcid èpsilon-aminocaproic i l'àcid tranaxèmic. Aquests fàrmacs actuen

inhibint la fibrinòlisis i consegüentment, estabilitzant el coàgul de fibrina. Per aquest motiu, tenen un risc potencial de desenvolupar trombosis en pacients amb una història familiar de tromboembòlia.18 Tanmateix, està contraindicat en pacients que presenten hemorràgies del tracte urinari ja que els coàguls que no lisen podrien causar obstrucció uretral. S'administren per via oral, intravenosa o tòpica.

- **Agents tòpics:** són preparacions que s'apliquen a als llocs d'hemorràgia exposats i actuen promovent l'hemostàsia i la cicatrització de les ferides.

- **Tractament hormonal:** és utilitzat pel tractament de la menorràgia en dones diagnosticades amb VWD i inclou els anticonceptius orals, preparacions d'estrogen-progesterona, o sistemes intrauterins amb alliberament de levonorgestrel.[68] En ocasions els sagnats severs durant la menstruació desencadenen una anèmia i, en aquests casos, l'administració de ferro pot derivar en una millor qualitat de vida.

En base als tractaments disponibles actualment, a la Taula 2 es mostren els més utilitzats depenent del tipus de VWD.[68]

Taula 2. Tractament d'elecció segons el tipus de VWD.

Tipus VWD		Primera opció*	Alternativa
Tipus 1 VWD		DDAVP	Concentrats de VWF/FVIII
Tipus 3 VWD	- inhibidor	Concentrats de VWF/FVIII	Concentrats de plaquetes
	+ inhibidor	rFVIII	rVIIa
Tipus 2A VWD		Concentrats de VWF/FVIII	DDAVP
Tipus 2M VWD		Concentrats de VWF/FVIII	DDAVP
Tipus 2B VWD		Concentrats de VWF/FVIII	-
Tipus 2N VWD		Concentrats de VWF/FVIII	DDAVP

** Es recomana la teràpia coadjuvant amb antifibrinolítics juntament amb les teràpies de primera opció en tots els tipus de VWD.*[68]

2.4.3 Noves opcions terapèutiques i perspectives de futur

Al 2016 va irrompre un nou producte pel tractament personalitzat de la VWD, el VWF recombinant (rVWF, Vonicog alfa). En comparació als concentrats de VWF/FVIII disponibles, el rVWF minimitza el risc de transmissió de patògens virals i prions derivats de la sang. Els assajos clínics de Fase I i Fase III han determinat que aquest producte no desencadena immunogenicitat, i és segur i eficaç pel tractament de la VWD. [83, 84] En conseüència, el rVWF ha estat aprovat per l'Administració d'aliments i medicaments (FDA), sota el nom de VONVENDI (Shire), pel tractament a demanda i control de sagnats en adults majors de 18 anys amb VWD.[80]

Respecte a altres possibles tractaments futurs per a la VWD el més destacat és la teràpia gènica. No obstant, aquest tractament tot i que té molt sentit en el cas de les hemofílies, només estaria indicada en aquells casos de VWD que presenten una simptomatologia molt greu i necessiten múltiples transfusions al llarg de la vida, cas infreqüent en la VWD.[85] Malgrat que encara no s'han realitzat assajos de teràpia gènica en humans, varies investigacions han demostrat l'eficàcia d'aquesta aproximació en la recuperació de l'expressió del VWF funcional en models animals.[86, 87]

3 El factor von Willebrand

3.1 El gen del VWF

El gen del VWF es localitza en l'extrem telomèric del braç curt del cromosoma 12, en la posició 12p13.31. És un dels gens més grans i complexos que codifiquen per proteïnes hemostàtiques descrit en humans amb una longitud de 178 kilobases (kb). Està format per un total de 52 exons d'entre 40 i 1.400 parells de bases (pb), sent l'exó 28 el més extens, i 51 introns d'entre 97 bp a 19,9 kb de longitud (Figura 7).[88, 89] El *VWF* es transcriu en un mRNA de 8,8 kb que donarà lloc a la proteïna del VWF.

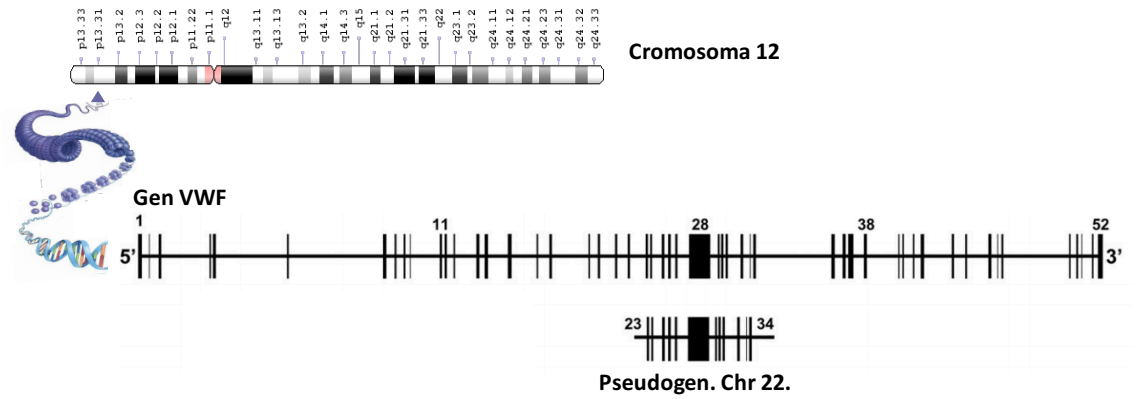


Figura 7. Localització i* organització *genòmica del VWF i del seu pseudogèn. *Chr indica cromosoma. Imatge adaptada de Baronciani et al.*[89]

El *VWF* és altament polimòrfic amb més de 100 polimorfisme d'un únic nucleòtid (SNPs) descrits en regió codificant segons la base de dades d'SNPs (http://www.ncbi.nlm.nih.gov/SNP). Addicionalment, al cromosoma 22 (22q11-13) existeix un pseudogèn (*VWFP*) parcial amb una longitud d'entre 21 i 29 kb, homòleg en un 97% a la regió que va des de l'exó 23 al 34 del *VWF*.[90] El fet que només hi hagi una divergència de seqüència del 3% en l'ADN suggereix un origen relativament recent (19-29 milions d'anys) del pseudogèn per duplicació parcial de gen. De fet, aquest pseudogèn es troba en humans i grans simis (bonobo, ximpanzé, goril·la i orangutan) però no en primats més llunyans.[88]

3.2 La proteïna del VWF

3.2.1 Estructura i dominis del VWF

El VWF és una glicoproteïna plasmàtica sintetitzada com a precursor proteic anomenat pre-pro-VWF amb una grandària de 2.813 aminoàcids (aa). Aquest precursor està compost per un pèptid senyal de 22 aa (prepèptid), un propèptid de 741 aa i la subunitat proteica madura de 2050 aa.[91] El VWF és molt ric en cisteïnes, 234 (8,3%) dels 2.813 residus.[92] L'anàlisi inicial del VWF va revelar que l'arquitectura molecular del pre-pro-VWF s'organitzava en 5 dominis homòlegs.[93] No obstant, gràcies a la utilització de la informació actualitzada sobre l'estructura d'altres proteïnes que contenen dominis

homòlegs, en combinació amb les tècniques de microscòpia electrònica d'última generació, l'arquitectura dels dominis del VWF ha estat redefinida recentment segons la següent estructura: D1-D2-D'-D3-A1-A2-A3-D4-C1-C2-C3-C4-C6-C6 (Figura 8).[94-96] Els canvis més destacats han estat la substitució dels dominis B i C, per sis dominis C consecutius i el descobriment de l'estructura interna dels dominis D. Els dominis D1 i

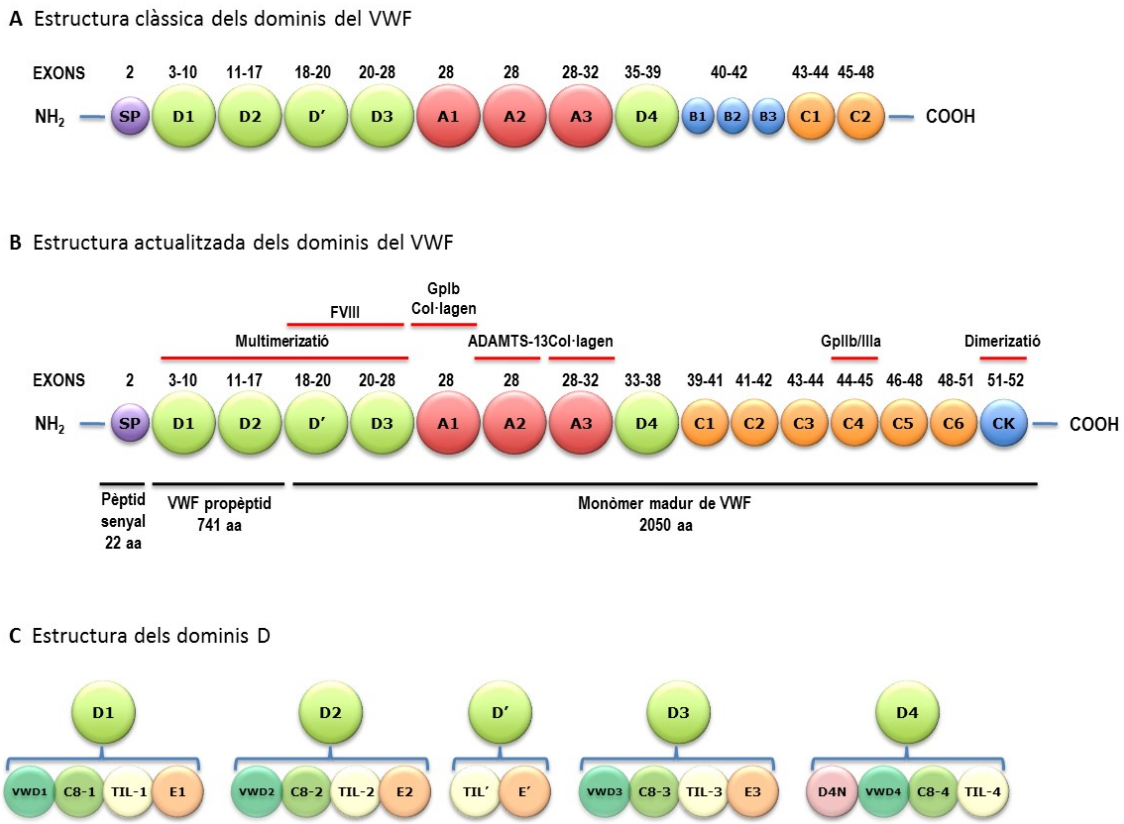


Figura 8. Representació esquemàtica de l'organització dels dominis segons l'estructura clàssica i remodelada de la proteïna VWF. *L'arquitectura molecular del VWF es caracteritza per la seva organització en dominis. A) Estructura clàssica dels dominis de VWF i els exons corresponents. B) Reorganització actualitzada dels dominis del VWF on s'indiquen els exons i les funcions corresponents de cada domini. C) Els dominis D estan formats per varies estructures independents. Les regions D1, D2 i D3 contenen un domini VW, una cisteïna-8 (C8), una estructura similar als inhibidors de tripsina (TIL) i un mòdul E. La regió D' no presenta el domini VW ni el C8. La regió D4 enlloc de presentar el mòdul E, té una seqüència única anomenada D4N. Imatge adaptada de Rauch et al.* [96]

D2 es localitzen al propèptid i la resta del D'al CK a la subunitat madura del VWF.

La majoria dels dominis del VWF contenen llocs d'unió funcionals que permeten la unió del VWF a altres molècules. En concret els dominis D' i D3 s'uneixen al FVIII i els dominis A1 i C4 a les plaquetes a través dels receptors GPIbα i $\alpha_2\beta_3$, respectivament. Tanmateix,

el domini A1 participa en la unió del VWF als col·làgens IV/VI, i conjuntament amb el A3 als col·làgens I/III (Figura 8B).[97-99] Per altra banda, els dominis del VWF intervenen en la biosíntesi del factor i presenten llocs d'escissió reconeguts per enzims implicats en la seva maduració. En aquest sentit, la furina actúa entre els dominis D2 i D' (Arg763-Ser764) provocant l'escissió del propèptid i la proteasa ADAMTS13 al domini A2 (Tyr1605-Met 1606) produint una gran varietat de multímers de diferents mides al torrent sanguini (veure secció 3.2.2).[17]

3.2.2 Biosíntesi, emmagatzematge i secreció del VWF

El VWF és sintetitzat com a pre-pro-VWF a les cèl·lules endotelials i als megacariòcits i està sotmès a un procés intracel·lular molt complex.[100] Al reticle endoplasmàtic, es produeix l'escissió del pèptid senyal, que donarà lloc al pro-VWF, i la formació de dímers, on dos subunitats s'uneixen de forma covalent mitjançant ponts disulfur en la regió carboxi-terminal. Els dímers pro-VWF migren cap a l'aparell de Golgi on el pH àcid i l'alta concentració de Ca^{2+} estimula la formació de multímers. Concretament, durant el procés de multimerització, els dominis D1-D' intervenen en l'alineament correcte dels dímers, seguidament el propèptid catalitza la formació dels ponts disulfur entre els dominis D3 (regió N-terminal) gràcies a la seva activitat de proteïna disulfur isomerasa.[100] El propèptid (D1-D2) s'escindeix per l'acció de la furina, tot i que continuarà unit de forma no covalent a la subunitat madura fins a la secreció de la proteïna. Durant tot el procés de biosíntesi, el VWF està sotmès a modificacions postraduccionals on la proteïna és glicosilada. El VWF sintetitzat s'emmagatzema en els cossos de Weibel-Palade (WPB, *weibel-palade bodies*) de les cèl·lules endotelials i als gràNuls α dels megacariòcits. Ambdues cèl·lules sintetitzadores el secreten al plasma i/o subendoteli de forma regulada en resposta a diversos estímuls, però únicament les cèl·lules endotelials també ho fan de forma constitutiva. Un cop a la circulació, el VWF presenta estructures multimèriques molt grans (UL-MWM, *ultra large molecular weight multimers*) que seran proteolitzades per l'ADAMTS13, donant lloc a una gran varietat de multímers de diferents mides, des de 500 a 20.000 kDa, que circulan al plasma (Figura 9).[92, 101]

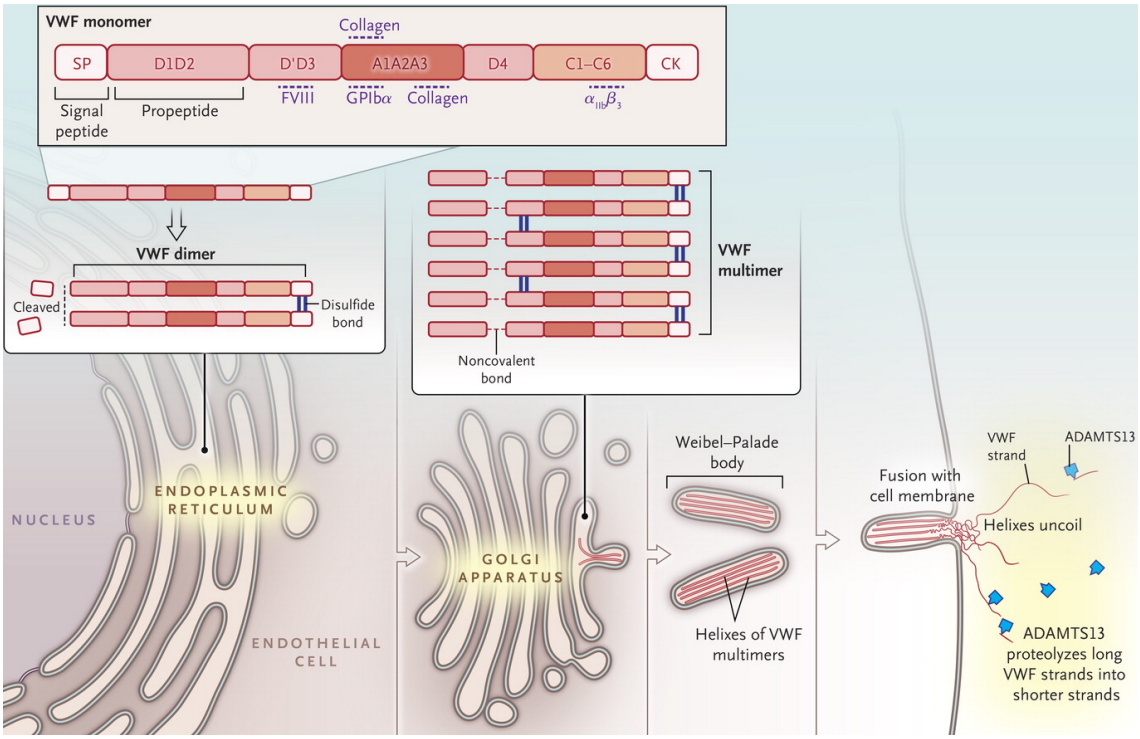


Figura 9. Representació esquemàtica del procés de biosíntesi del VWF en una cèl·lula endotelial. *Imatge de Leebeek FW et al.*[100]

El VWF té una vida mitja d'entre 8 i 12h a la circulació i, al igual que altres proteïnes plasmàtiques, és sensible a canvis físics, com l'oxidació i la proteòlisi, que alteren la seva funcionalitat. Tot i que encara està en vies d'investigació, gràcies a estudis realitzats en models murins i cultius cel·lulars, s'ha reportat que el VWF podria ser eliminat de la circulació per les cèl·lules endotelials[102] o pels macròfags que actuarien al fetge i a la melsa.[103, 104]

3.3 Paper del VWF en l'hemostàsia

L'hemostàsia és el conjunt de mecanismes fisiològics de l'organisme encarregats de mantenir la integritat i permeabilitat del sistema circulatori que permet aturar els sagnats després d'una lesió vascular. Davant d'una hemorràgia es produeix una vasoconstricció que provoca una reducció del flux sanguini al lloc de la lesió. Seguidament, es produeix la formació del tampó hemostàtic plaquetar inestable (hemostàsia primària), que posteriorment, s'estabilitzarà gràcies a un component

anomenat fibrina (hemostàsia secundària). En aquest context, el VWF desenvolupa un paper fonamental tant en l'hemostàsia primària com en la secundària.

3.3.1 El VWF a l'hemostàsia primària

L'hemostàsia primària consisteix en l'acumulació de les plaquetes al lloc on s'ha produït una lesió vascular, que donarà com a resultat la formació d'un tampó blanc o plaquetari, inestable i fràgil, que actuarà aturant l'hemorràgia.

En absència d'alteracions hemostàtiques, el VWF circula pel plasma amb una conformació globular. Davant d'una lesió vascular, el VWF reconeix el col·lagen subendotelial que queda exposat i, alhora, interacciona amb altres regions del vas sanguini com els proteoglicans de la matriu cel·lular gràcies als dominis A1, D' i D3 on es localitzen els llocs d'unió a heparina.[105, 106] Una vegada el VWF és immobilitzat a la superfície cel·lular, aquest canvia de conformació i es desenrotlla formant una cadena molt més adhesiva (Figura 10). En aquest punt, el VWF promou **l'adhesió de les plaquetes** al subendoteli mitjançant la interacció amb el receptor GpIbα (del complex GPIb-IX-V). Aquest primer contacte provoca una reducció de la velocitat d'aquestes cèl·lules al flux sanguini, facilitant la seva unió al col·lagen subendotelial, el que es tradueix en una adhesió més estable del tampó. Consecutivament les plaquetes s'activen i alliberen al lloc de la lesió el contingut emmagatzemat als gràmuls α (VWF i altres substàncies). Finalment es produeix **l'agregació plaquetària** mitjançant la

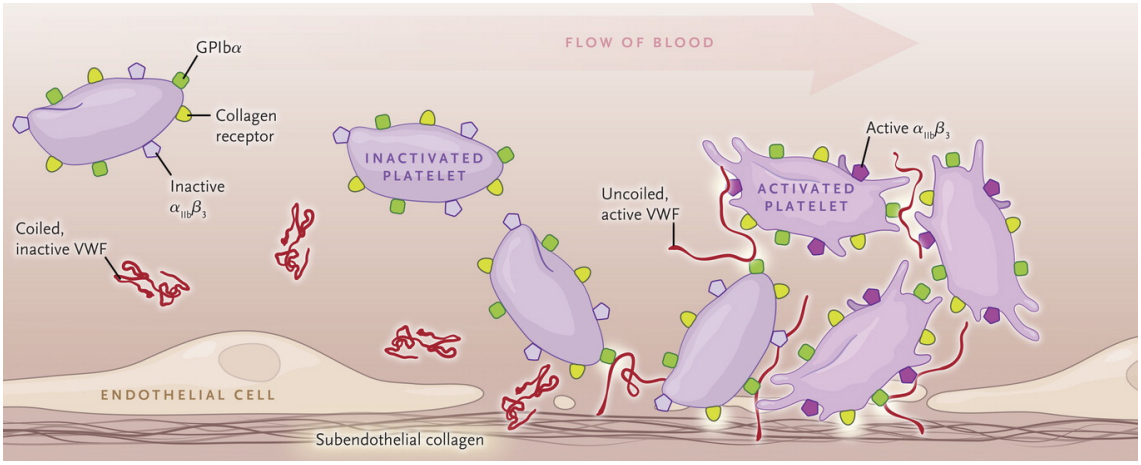


Figura 10. Representació esquemàtica del paper del VWF en l'activació i agregació plaquetària durant l'hemostàsia primària. *Imatge de Leebeek FW et al.*[100]

interacció del VWF amb la integrina αIIbβ3 (del complex GPIIb-IIIa) situada a la membrana de les plaquetes (Figura 10).[107]

3.3.2 El VWF a l'hemostàsia secundària

En l'hemostàsia secundària es duen a terme reaccions complexes entre proenzims, enzims i cofactors que tenen com a resultat final la formació de fibrina. Segons el model clàssic, establert l'any 1964, es considera que el procés de coagulació és una cascada enzimàtica seqüencial, on l'activació d'un factor activa al següent.[108, 109] En base a aquest model, existeixen dues vies d'activació: la via extrínseca iniciada pel factor tissular (FT) i FVII, i la via intrínseca iniciada pel FXII. Ambdues convergeixen en una via comú per activar al FX i continuar fins a la formació de la trombina que permetrà la conversió del fibrinogen a fibrina (Figura 11).

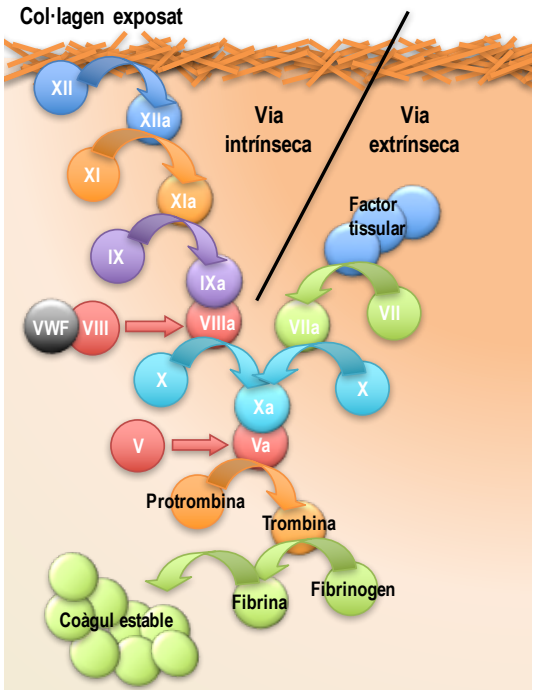


Figura 11. Representació esquemàtica del model clàssic de l'hemostàsia secundària.

Investigacions posteriors han canviat el paradigma de l'hemostàsia secundària ja que s'ha demostrat que ambdues cascades, intrínseca i extrínseca, no actuen de forma independent i que les plaquetes participen activament en la coagulació. Tots aquests canvis són els que es recullen al model cel·lular de la coagulació (Figura 12).[110, 111] En concret, aquest model es caracteritza per tres fases seqüencials: iniciació, amplificació i propagació. En la fase inicial, com a conseqüència d'una lesió vascular, el FT queda exposat al torrent sanguini i s'uneix al FVII afavorint la seva activació (FVIIa). El complex FT/FVIIa permet l'activació posterior dels factors IX i X. El factor Xa genera petites quantitats de trombina (FIIa) a nivell local. A continuació, durant la fase d'amplificació,

les quantitats de FIIa acabaran d'activar a les plaquetes que s'han adherit al lloc de la lesió. Paral·lelament, la FIIa produeix l'activació dels factors FXIa, el FVa (cofactor al factor X) i FVIIIa (cofactor al factor IX), el que augmenta exponencialment l'activitat catalítica dels factors IX i X. A l'última fase, el FXIa converteix el FIX en FIXa i el complex FIXa/FVIIIa catalitza la conversió de FX a FXa. Finalment, el complex FXa/FVa produeix quantitats suficients de trombina per formar fibres de fibrina massivament i estabilitzar el coàgul. La primera fase té lloc a les cèl·lules que expressen el FT, mentre que les dues últimes fases tenen lloc a la superfície de les plaquetes activades.

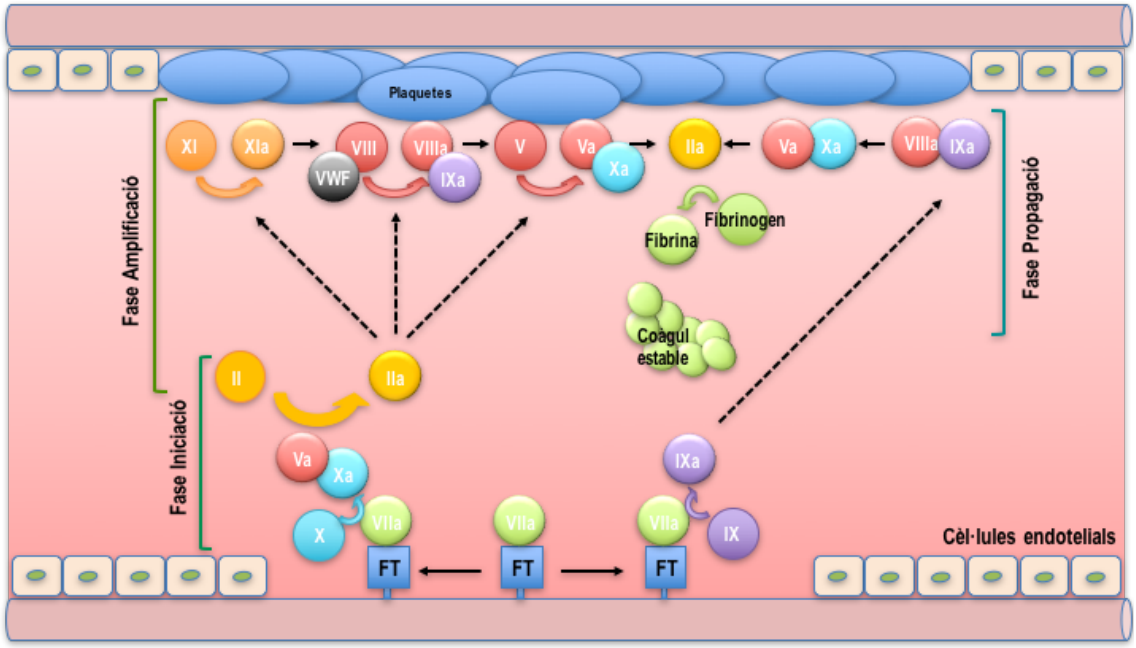


Figura 12. Representació esquemàtica del model cel·lular de l'hemostàsia secundària*. FT indica factor tissular.*

En aquest escenari, el VWF actua **estabilitzant el FVIII** a la circulació, evitant la seva degradació proteolítica prematura. Cada monòmer de VWF s'uneix a una molècula de FVIII, prolongant la seva vida mitja i localitzant-lo eficientment en el lloc de lesions vasculars. El complex VWF-FVIII s'uneix de forma no covalent entre els dominis D' i D3 del VWF (Figura 8) i els aminoàcids 1669-1689 del FVIII. Durant la coagulació, el FVIII és activat per la trombina que provoca l'escissió al residu Arg1689 i, com a conseqüència, la destrucció del lloc d'unió al VWF i l'alliberament del FVIII activat (FVIIIa).[112]

3.4 Paper del VWF en altres processos patològics de la biologia vascular

El VWF té un caràcter molt versàtil a nivell funcional gràcies als múltiples dominis de la proteïna. En conseqüència, a banda del paper àmpliament descrit en l'hemostàsia, s'han reportat diferents funcions[95] d'aquest factor relacionades freqüentment amb processos patològics de la biologia vascular:

- **Actua potenciant la inflamació**: la inflamació forma part de la reacció autoimmune de l'organisme envers d'una agressió externa (lesió, infecció o traumatisme). Aquesta té com a objectiu final reclutar leucòcits per aïllar i eliminar l'agent patogen, així com, reparar el teixit danyat. En aquest context, el VWF actua com a marcador del procés inflamatori ja que les cèl·lules endotelials s'activen i alliberen el contingut emmagatzemat als WPB, sent el VWF un dels components majoritaris.[113] Per altra banda, potència el reclutament dels leucòcits al lloc de la inflamació per via directa, unint-se als receptors leucocitaris integrina β2 i PSGL1 (*P-selectin glycoprotein ligand 1*) i per via indirecta, garantint la correcta expressió del receptor P-selectina a la superfície de les cèl·lules endotelials, necessari pel reclutament dels leucòcits al lloc de la lesió.[114-116]

- **Regula negativament l'angiogènesi**: aquest procés consisteix en la formació de vasos sanguinis nous. Estudis *in vitro* i *in vivo* han demostrat que el VWF regula negativament la proliferació de les cèl·lules endotelials ja que una reducció d'aquest factor provoca un augment de l'expressió del VEGFR2 (*Vascular Endothelial Growth Factor Receptor 2*) en les cèl·lules endotelials.[117] Malgrat no es coneix el mecanisme d'acció, es postula que el VWF actua gràcies a la seva interacció amb varies molècules pro-angiogèniques (Angiopeptina2, Galactina-1 i Galactina-2).[118] Aquest efecte s'evidencia pel fet que un 10% dels pacients amb VWD i, per tant, deficients pel VWF presenten angiodisplasia, una malformació vascular del colon que provoca el sagnat del tracte gastrointestinal.[119]

- **Contribueix a la proliferació cel·lular de les cèl·lules musculars llises**: després de lesions vasculars, el VWF pot penetrar a la capa íntima dels vasos, on contacta amb les cèl·lules musculars llises.[120] La acumulació del VWF en la capa intima s'ha associat amb l'engrossiment d'aquesta regió, suggerint la contribució d'aquest factor en la proliferació de cèl·lules musculars llises.[121]

- **Indueix l'apoptosi cel·lular:** la capacitat del VWF per induir la mort cel·lular s'ha demostrat dirigida: a les plaquetes, mitjançant interaccions amb el receptor plaquetar GpIbα,[122] i; a les cèl·lules tumorals, via interaccions amb la integrina $\alpha V\beta 3$.[123, 124]

En general, les funcions del VWF no relacionades amb l'hemostàsia s'han descrit gràcies als estudis *in vitro* o *in vivo* amb l'ús de models animals amb VWD. En aquest sentit, és destacable que la majoria de les funcions reportades es troben en vies d'investigació i encara són necessàries més evidències per determinar el mecanisme d'acció específic del VWF en cada un d'aquests processos patogènics.

4 Diagnòstic genètic de la malaltia de von Willebrand

4.1 Transcendència del diagnòstic genètic en la VWD

L'anàlisi genètic de la VWD esdevé essencial en determinades situacions per a establir un diagnòstic precís i definitiu. En aquest sentit, l'estudi genètic en la VWD de tipus 2 es especialment útil en la pràctica clínica quan no es disposa de proves especialitzades com el RIPA (per discriminar entre els tipus 2A i 2B), l'anàlisi de multímers (per discriminar entre els tipus 2A i 2M) o el VWF:FVIIIB (per discriminar el tipus 2N de l'hemofília A).[125] Tanmateix, permet la diferenciació entre tipus 1 greu i tipus 3, a vegades difícils de diferenciar tenint en compte únicament les proves de laboratori. A més a més, permet establir la distinció de forma definitiva de la VWD en front de les seves genocòpies, patologies amb clínica similar però causa genètica diferent. Un exemple són la VWD de tipus 2N i l'hemofília A causada per mutacions al *F8*, i la VWD

de tipus 2B i el pseudo-VWD (PT-VWD, *platelet type*) causat per mutacions a l'exó 2 del gen GP1BA.[126]

Per altra banda, el diagnòstic molecular també te implicacions en la decisió de la pauta de tractament. En aquest sentit la identificació de mutacions causals de tipus 2B són una evidència per no utilitzar DDAVP com a tractament, ja que està contraindicat en aquest subtipus.[127]

L'estudi genètic en la VWD de tipus 3 és beneficiós en la predicció del risc de desenvolupament d'anticossos com a resultat del tractament amb VWF exògen. Els anticossos són més freqüents en pacients que presenten delecions genètiques parcials o totals.[44, 128] Addicionalment, en aquesta patologia la seva utilització té especial interès a l'hora de realitzar el diagnòstic prenatal a partir de vellositats coriòniques o aminòcits, en aquelles famílies que ja han tingut un fill diagnosticat amb aquest tipus de patologia, o bé, es coneix que els progenitors presenten VWD.[125]

Per establir el diagnòstic genètic es disposa de dues aproximacions: els mètodes indirectes, cada vegada menys utilitzats, i els mètodes directes.

4.2 Mètodes indirectes

Es basen en estudis de lligament genètic utilitzant marcadors polimòrfics intra- o extragènics. Aquest marcadors són variacions de la seqüència del genoma dels individus que permeten establir un seguiment del cromosoma portador del gen defectuós, però no identifiquen la mutació patogènica. Existeixen varis tipus de polimorfismes al genoma utilitzats en diagnòstic molecular: aquells que impliquen el canvi d'un sol nucleòtid al DNA anomenats SNPs (de l'anglès *Single Nucleotide Polymorphism*), o aquells que suposen la repetició de seqüències de varis nucleòtids en tàndem anomenats VNTR (de l'anglès *Variable Number of Tandem Repeats*). Es coneixen dos tipus de VNTR depenent del número de nucleòtids que es repeteixen: els minisatèl·lits (6-100bp) i els microsatèl·lits (2-5bp), també conegudes com a STRs (de l'anglès *Short Tàndem Repeats*). L'anàlisi indirecta es una aproximació diagnòstica ràpida i senzilla.

No obstant, presenta alguns inconvenients com la necessitat d'estudiar diferents familiars per establir el cromosoma portador del gen defectuós o la presència de polimorfismes no informatius. Addicionalment, la recombinació genètica o el mosaïcisme poden dificultar l'obtenció de resultats concloents.

En el cas de la VWD, els marcadors polimòrfics més documentats i àmpliament acceptats pel diagnòstic molecular són 4 STRs.[129] Tres d'aquestes són repeticions de tertanucleòtids (ATCT) localitzades a l'intró 41 que es coneixen com a VNTR1, VNTR2 i VNTR3.[130, 131] L'altre marcador utilitzat és la repetició d'un dinucleòtid (GT) localitzat a la regió promotora del *VWF* (VWP) (Figura 13). Aquesta tècnica es pot utilitzar en consell genètic i diagnòstic prenatal de la VWD, per fer estudis d'associació entre el gen i el desenvolupament de la patologia i per investigar el possible origen comú d'una mutació identificada en varies famílies. [132, 133]

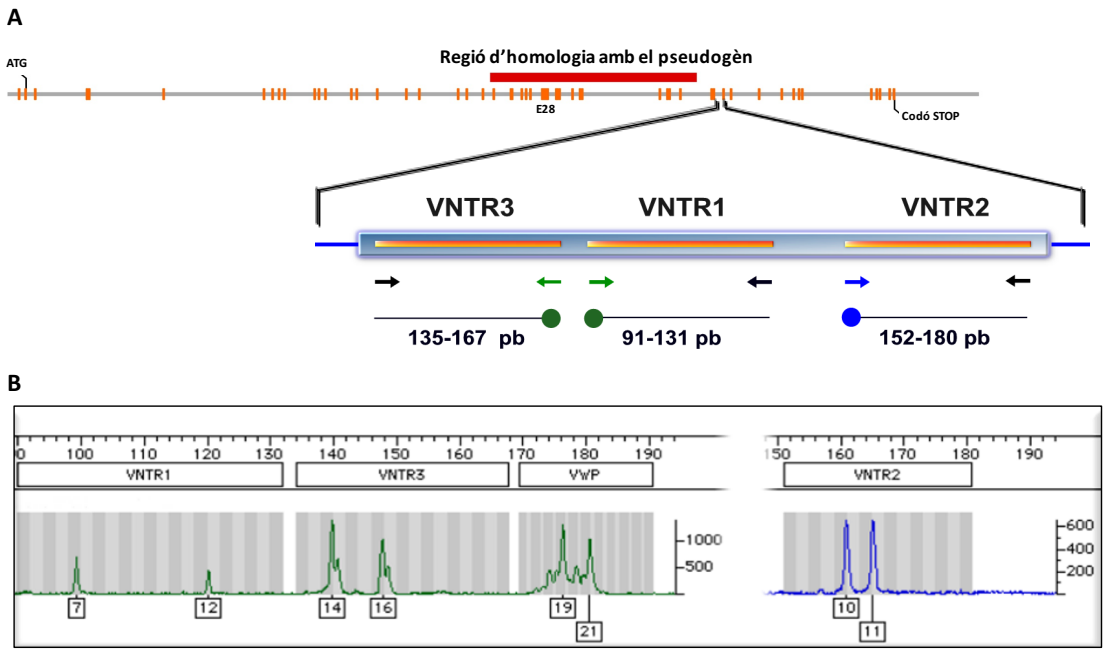


Figura 13. Localització de les STRs de l'intró 41 i visualització dels resultats juntament amb la STR del promotor. *A. Es mostren les VNTR-1, VNTR-2 i VNTR-3 de l'intró 41 i la mida esperada dels productes de PCR. Els primers es van etiquetar amb un fluorocrom (TET, verd, o FAM, blau) tenint en compte la sobreposició de la mida dels amplímers. B. Visualització dels resultats d'una mostra representativa analitzats amb el programa Genotyper (Applied Biosystems). Es mostra l'al·lel identificat per a cada polimorfisme de les 4 STRs. Imatge adaptada de Vidal et al.*[129]

4.3 Mètodes directes

Aquests mètodes tenen la finalitat de determinar el defecte genètic responsable de la patologia familiar. A grans trets, existeixen dues aproximacions pel diagnòstic directe: mitjançant procediments de cribratge de mutacions, o bé; mitjançant la seqüenciació directa del gen d'interès. Aquesta última, és la més simple conceptualment ja que té com a objectiu la lectura de la seqüència de nucleòtids de les regions del DNA implicades en l'expressió de la proteïna, permetent la identificació de la mutació responsable de la patologia. Actualment, existeixen dues alternatives metodològiques: la seqüenciació tradicional de *Sanger* i la seqüenciació massiva o *Next Generation Sequencing* (NGS).

4.3.1 Cribratge de mutacions

Tenen la finalitat d'analitzar regions o segments amplificats del gen de manera que alguna característica fisica i/oquímica permeti diferenciar entre un fragment control i el seu homòleg contenint una alteració puntual. S'han descrit diferents tècniques de cribatge, utilitzades històricament per a la detecció de mutacions al *VWF*, com *Single Strand Conformational Polymorphism* (SSCP), *Conformational Sensitive Gel Electrophoresis* (CSGE), *Denaturing High Performance Liquid Chromatography* (dHPLC) i *Denaturing gradient gel electrophoresis* (DGGE).[134-137] Totes elles tenen la peculiaritat que després d'analitzar els diferents fragments i determinar els sospitosos de ser portadors d'una mutació, aquests han de ser seqüenciats per corroborar l'existència de l'alteració i si aquesta podria ser considerada la responsable de la malaltia. No obstant, actualment aquestes tècniques estan en desús ja que impliquen un esforç considerable de manipulació per part de personal altament qualificat i la naturalesa altament polimòrfica del *VWF* té un efecte negatiu en la seva eficiència.

4.3.2 Seqüenciació tradicional de *Sanger*

Al 1977, Frederick Sanger va desenvolupar un mètode per seqüenciar el DNA, conegut com a seqüenciació amb terminadors dideoxinucleòtids o mètode de *Sanger*,[138] pel qual

va ser guardonat amb el premi Nobel al 1980. Aquest mètode consisteix en sintetitzar de forma seqüencial la cadena complementària del fragment de DNA d'interès gràcies a la utilització de la polimerasa de DNA, quatre deoxinucleòtids (dNTPs) i quatre dideoxinucleòtids (ddNTPs) que no presenten el grup hidroxil del carboni 3' provocant l'aturada de l'elongació de la cadena. En versions actuals i automatitzades del mètode, són necessàries moltes còpies del fragment de DNA d'interès que, en general, s'obtenen per amplificació mitjançant PCR. A continuació, el DNA es desnaturalitza permetent la unió dels oligonucleòtids per complementarietat de bases a les regions flanquejants del DNA d'interès. La unió aleatòria d'un ddNTP (marcats amb diferents fluorocroms) a la cadena creixent evitarà l'addició d'un nou nucleòtid per part de la polimerasa i consegüentment la parada de l'elongació de la cadena de forma no reversible, donant lloc a molècules de diferents longituds marcades fluorescentment. Aquestes molècules resultants són separades per mida mitjançant electroforesi capil·lar. Els fragments migren pel capil·lar i, en arribar al làser, es produeix l'emissió de la llum del fluoròfor que serà detectada per una càmera permetent la identificació del nucleòtid incorporat.[139] D'aquesta manera es determina la seqüència ordenada de nucleòtids del fragment d'interès que s'interpreta en cromatogrames i es compara amb una seqüència consens per determinar la presència de mutacions.

- **Seqüenciació de *Sanger* aplicada al diagnòstic molecular de la VWD**

Durant anys l'estudi molecular de la VWD mitjançant la tecnologia de *Sanger* s'ha limitat a la investigació bàsica i la seva aplicació en la pràctica clínica habitual s'ha retardat significativament a causa de la grandària i complexitat del gen.[140] Tanmateix, la presència d' un pseudogèn parcial al cromosoma 22 i l'elevat nombre de polimorfismes al *VWF* impliquen la necessitat d'un disseny d'oligonucleòtids específics que permetin amplificar de forma exclusiva el *VWF*, així com evitar l'amplificació preferencial d'un al·lel en concret.[141] La seqüenciació de *Sanger* és considerada una eina relativament costosa econòmicament per utilitzar-se de forma generalitzada en la seqüenciació completa del *VWF* de tots els pacients. Per aquest motiu la majoria de laboratoris de rutina s'han limitat a la seqüenciació d'exons específics segons el tipus de VWD.[133] En aquest sentit, l'exó 28 del *VWF* és seqüenciat davant la sospita de 2A, 2B o 2M VWD.

En el cas del tipus 2A VWD si no s'ha identificat mutació a l'exó 28, l'estudi genètic s'amplia als exons del 11 al 17 i del 50 al 52. Els exons del 18 al 20 són analitzats pel tipus 2N VWD. Pels tipus de VWD quantitatius, donat que les mutacions responsables de la patologia poden estar repartides al llarg del gen, es requereix la seqüenciació completa del *VWF.*[132] En aquest sentit, el nostre grup va dissenyar i optimitzar l'any 2019 un procediment ràpid i senzill per a la seqüenciació del *VWF* complet que va ser implementat com a mètode de rutina per a tots els tipus de VWD (Figura 14).[142]

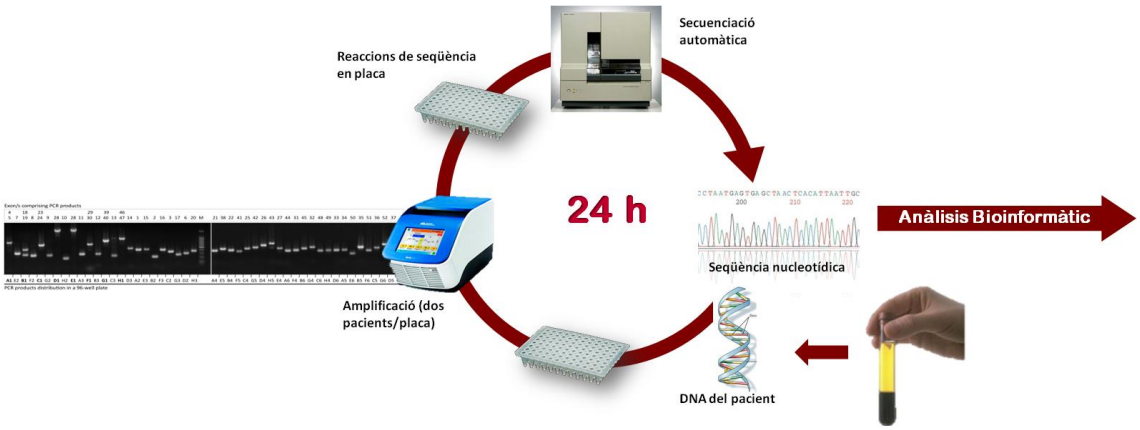


Figura 14. Representació esquemàtica del procediment de diagnòstic molecular del VWF basat en la tècnica de Sanger. *A partir de la sang del pacient s'obté el DNA i s'utilitza com a motlle per amplificar les regions d'interès del VWF mitjançant 52 reaccions de PCR. A continuació, es realitzen les reaccions de seqüència i mitjançant programes bioinformàtics es determina l'ordre dels nucleòtids de cada amplicó.*

4.3.3 Next Generation Sequencing

La introducció de la NGS ha provocat un canvi substancial en la capacitat de les tècniques de seqüenciació del DNA i ha revolucionat tant la investigació bàsica en genètica com la seva aplicació clínica. Aquesta tecnologia és substancialment més ràpida i econòmica que la seqüenciació tradicional de *Sanger* i es basa en la seqüenciació simultània i en paral·lel de milers o milions de molècules de DNA. Un exemple paradigmàtic va ser la seqüenciació del primer genoma humà complet mitjançant NGS (any 2008): en només quatre mesos i mig, a un cost de <1,5 milions de dolars,[143] mentre que al Projecte Genoma Humà, abordat mitjançant la tecnologia tradicional, va requerir 13 anys i a un cost de 2.700 milions de dòlars.[144]

- **Plataformes de segona generació**

Les plataformes de NGS van irrompre comercialment al mercat al 2005 amb la presentació del primer seqüenciador anomenat *454 Genome Sequencer* (*454 Life Sciences*, empresa adquirida al 2007 per Roche). Actualment, es disposa de varies plataformes de seqüenciació al mercat comercialitzades per Illumina i Thermofisher Scientific.[145] Aquestes es basen en diverses estratègies però totes comparteixen un denominador comú: l'amplificació clonal del DNA abans de procedir a la seqüenciació. Per contra, es diferencien en la química utilitzada, en la longitud de lectura, en el nombre de molècules que es poden seqüenciar en paral·lel i en la taxa d'error.[146, 147] En general, una de les principals limitacions que presenten aquestes plataformes és la curta longitud de les seqüències que poden llegir (35-700 bp) que provoca dificultats en el procés d'acoblament *de novo* i en l'alineament enfront de seqüències de referència, en particular en les regions repetitives.[148] En l'actualitat, Illumina és la plataforma més àmpliament utilitzada ja que ha mostrat una elevada eficiència de seqüenciació.[145]

	iSeq 100	MiniSeq	MiSeq	NextSeq	HiSeq	HiSeq X	NovaSeq 6000
Run Time	9-17.5h	4-24h	4-55h	12-30h	<1-3.5 d* 7h-6 d†	<3days	16-36h
Output	1.2 Gb	7.5 Gb	15 Gb	120 Gb	1500 Gb	1800 Gb	6000 Gb
Read Length	2x150 bp	2x150 bp	2x300 bp	2x150 bp	1x150 bp	2x150 bp	2x150 bp

Figura 15. Familia de seqüenciadors d'Illumina. **HiSeq 3000 System i HiSeq 4000 System. † HiSeq 2500 System*

Aquesta disposa d'una gran varietat d'aparells disponibles, des de seqüenciadors de baix rendiment com el iSeq a altres de molt alt rendiment com NovaSeq 6000 (Figura 15).

La tecnologia de seqüenciació d'Illumina inclou 3 passos bàsics. En primer lloc, la preparació de llibreries de les mostres a seqüenciar. Durant aquest procés es generen fragments curts del DNA d'interès seguit de la unió dels adaptadors i de seqüències de nucleòtids específics per pacient (índex o codi de barres), que permet seqüenciar mostres de diferents individus simultàniament. En segon lloc, la generació de clústers. Els fragments són immobilitzats en una superfície sòlida anomenada *flow-cell* mitjançant el reconeixement dels adaptadors específics. Cada un d'aquest fragments s'amplifica de forma clonal en pont donant com a resultat al voltant d'un milió de còpies d'una molècula de DNA, el que es coneix com a clúster (Figura 16).[149] En tercer lloc, la seqüenciació mitjançant el mètode de seqüenciació per síntesi de tipus terminació cíclica reversible (SBS-CRT). Aquest mètode es basa en la utilització d'una DNA polimerasa i de terminadors reversibles fluorescents que prevenen l'elongació, marcats amb fluorescència i amb el grup 3'-OH bloquejat (dNTP). Durant cada cicle de

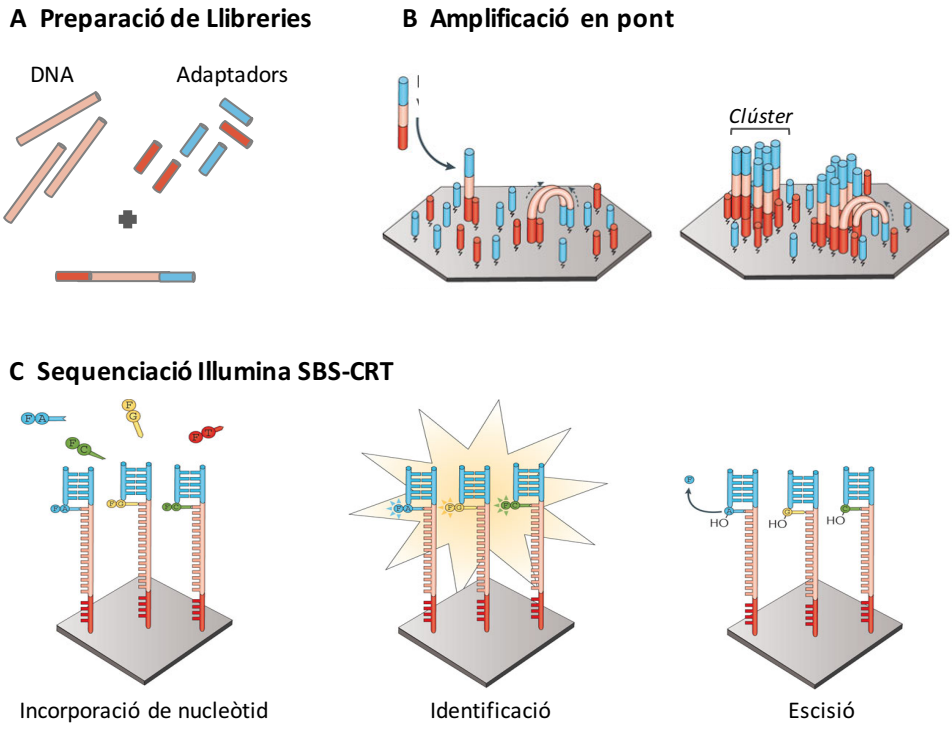


Figura 16. Representació esquemàtica del workflow de seqüenciació de la plataforma Illumina. *A. Durant la preparació de llibreries s'obtenen fragments curts amb els adaptadors. B. La flow-cell conté primers units covalentment a la seva superfície que són complementaris als adaptadors dels fragments permetent la immobilització d'aquests. A continuació, té lloc l'amplificació clonal per formar els clústers (100-200 milions de clústers). C. La seqüenciació dels fragments utilitza la tecnologia de seqüenciació per síntesi de terminació cíclica reversible. Imatge adaptada de Goodwin et al.*[146]

seqüenciació, s'afegeix un únic dNTP marcat a la cadena d'àcids nucleics. La base incorporada es detecta després de l'excitació làser que permet capturar la fluorescència emesa de cada clúster mitjançant la reflexió interna total de fluorescència (TIRF). Finalment, s'escindeix el fluorescent i el grup inhibidor permetent la regeneració del grup 3'-OH del nucleòtid i l'inici del següent cicle.

- **Plataformes de tercera generació (TGS)**

Aquestes plataformes, conegudes també com a *Next-Next Generation Sequencing*, es caracteritzen per la capacitat de seqüenciació d'una única molècula de DNA i la detecció de nucleòtids a temps real, sense la necessitat d'una amplificació prèvia.[150] Helicos Biosciences, va ser la primera empresa que va anunciar una tecnologia de seqüenciació d'aquestes característiques, tot i que les més destacades actualment són les empreses de Pacific BioScienes i la Oxford Nanopore Technologies (ONT). Pacific BioScienes va presentar un sistema basat en la seqüenciació del DNA en pous anomenats ZMWs (*Zero-Mode Waveguides*) que contenen una polimerasa immobilitzada al fons del pou

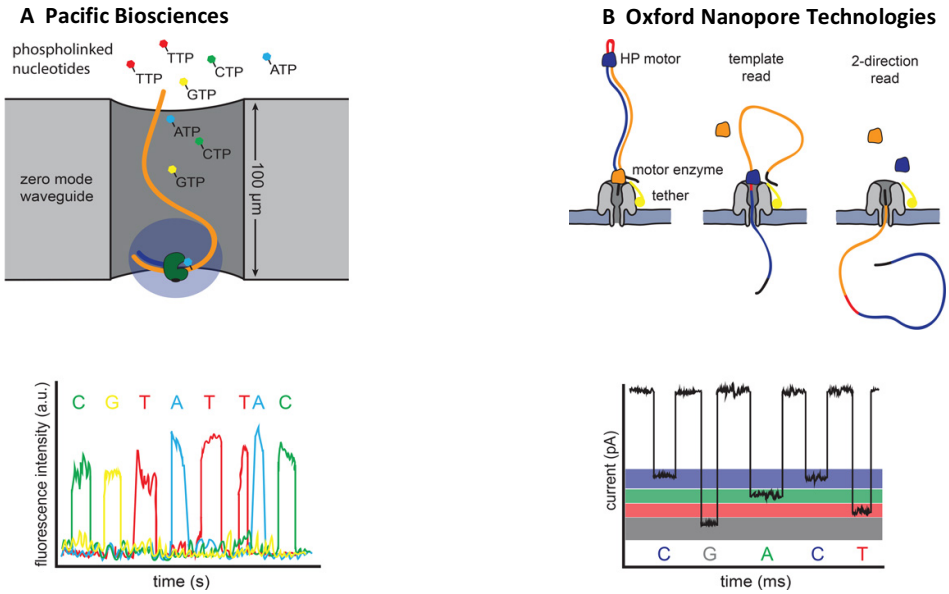


Figura 17. Il·lustració de les tecnologies de seqüenciació a temps real de les plataformes de tercera generació. *A. Seqüenciació SMRT de Pacific Biosciences. La polimerasa s'immobilitza al fons transparent de cada pou anomenat ZMW. Els quatre nucleòtids marcats amb fluorescència permeten la polimerització del motlle de DNA. La incorporació d'una base pausa momentàniament la polimerasa permetent la identificació de la base incorporada. B. Seqüenciació Nanopore de ONT. El DNA s'uneix a dos adaptadors que dirigeixen el DNA cap al porus. Els canvis de corrent induïts pel pas de nucleòtids a través seu s'utilitza per discriminar les bases. Imatge adaptada de Reuter et al.*[150]

(Figura 17).[151, 152] Uns anys més tard, l'empresa ONT va presentar un altre sistema de seqüenciació utilitzant la tecnologia nanopore que discrimina les bases de DNA a mesura que passen per un canal nanomètric on es detecta el canvi de corrent provocat per la transmissió de cada base.[153] Les seves plataformes més destacades són el GridION i el MinION, sent aquest últim el seqüenciador de mida més reduïda del mercat, similar a un dispositiu d'emmagatzematge portàtil USB. Les plataformes de seqüenciació de tercera generació a partir d'un sola molècula tenen la capacitat de realitzar lectures molt més llargues (10-100 kb) que les plataformes de segona generació. No obstant, encara presenten limitacions importants com l'elevada taxa d'error.[146]

5 Alteracions moleculars al VWF

5.1 Bases moleculars associades als diferents tipus de la VWD

Gràcies a l'estudi genètic en pacients amb VWD s'han identificat centenars de mutacions que han permès fonamentar la investigació i establir les singularitats existents en les bases moleculars dels diferents tipus de VWD. En aquest sentit, les mutacions es diferencien tant en la localització al llarg del gen com en el tipus de mutació més freqüent reportat com a causant de la patologia (Taula 3).

La **VWD de tipus 1** presenta una patró d'herència autosòmica dominant amb penetrància incompleta (una alteració genètica pot estar associada a diferents graus de severitat observats als pacients) i expressió variable. Les mutacions de tipus 1 VWD es troben distribuïdes al llarg de tot el *VWF* i les més freqüents són de tipus *missense*. Els mecanismes patogènics associats a aquestes mutacions són molt heterogenis. Aquests inclouen: l'alteració de la secreció del VWF degut a mutacions en la regió de la dimerització (p.ex p.Cys2739Tyr) o multimerització del VWF (p.ex p.Cys1130Phe i p.Cys1149Arg); [155] l'augment de la susceptibilitat a la proteòlisi mediada per l'ADAMTS13 (p.Tyr1584Cys);[156] l'augment de l'aclariment del VWF (p.Arg1205His, mutació Vicenza);[157] la reducció de la síntesi del factor degut a mutacions en els llocs

d'unió dels factors de transcripció, com la c.-1522_-1510del13[158] o; a variacions que afecten a l'*splicing* del mRNA.[159]

Taula 3. Característiques genètiques dels tipus de la VWD.

Tipus VWD	Patró d'herència	Localització de les mutacions	Tipus mutació	Mecanisme
1	AD/ co-dominant	1-52, promotor	*Nonsense, missense, indel*	Deficiència quantitativa parcial
3	AR	1-52	*Nonsense, missense, indel*	Deficiència quantitativa total
2A	AD/AR	2-28, 51, 52	*Missense*	Alteració en la multimerització
2B	AD	28	*Missense*	Augment de la unió a plaquetes
2M	AD	28-32	*Missense*	Disminució de la unió a plaquetes i col·làgens
2N	AR	17-27	*Missense, Null*	Disminució de la unió a FVIII

Mutacions "null" són aquelles que provoquen la no expressió de l'al·lel i inclouen les de tipus nonsense, splicing i insercions o delecions que provoquen un canvi en la pauta de lectura . AD indica autosòmica dominant; AR, autosòmica recessiva.[154]

La **VWD de tipus 3** presenta un patró d'herència autosòmic recessiu i, per aquest motiu, són necessàries dues mutacions en heterozigosi composta o una mutació en homozigosi per donar lloc a la manca d'expressió total del VWF. Les mutacions de tipus 3 VWD es localitzen al llarg de tot el gen i la majoria donen lloc a al·lels nuls. Les més freqüents són de tipus *nonsense*, *frameshift* o mutacions potencials d'*splicing* (PSSM) i, en menor freqüència, les conversions gèniques i les *missense*.[160-162] La conversió gènica consisteix en la recombinació homòloga entre el *VWF* i el seu pseudogèn, donant com a resultat la introducció de mutacions derivades de la seqüencia del *VWFP*, sent el canvi més freqüent la p.Gln1311Ter.[163, 164] El mecanisme patogènic de les mutacions de tipus 3 s'associa predominantment a la biosíntesi disminuïda i/o la interrupció del plegament de la proteïna que inhibeix la seva secreció.[160]

En la **VWD de tipus 2**, els subtipus 2A, 2B i 2M presenten un patró d'herència autosòmic dominant amb penetrància completa, a excepció d'alguns tipus 2A on es requereix la presència de mutacions en homozigosi o heterozigosi composta per desenvolupar la patologia.[17]

La VWD 2A està causada per mutacions principalment *missense* localitzades als dominis D1-D2, D3, A2-A1 i CK, sent el domini A2 on es localitzen el major número de mutacions amb un 82%, seguit dels dominis D2 i CK amb un 8% i, finalment, el domini D3 amb un 1%.[133] Els mecanismes patogènics descrits associats a les mutacions de tipus 2A inclouen: l'alteració de la síntesi de dímers i multímers del VWF, un augment de la proteòlisi dels HMWMs per la proteasa ADAMTS13 i la retenció intracel·lular del VWF.[165, 166] En tots els casos, aquestes alteracions tenen com a conseqüència la pèrdua dels HMWM. Tanmateix, algunes variants de tipus 2A poden influir simultàniament en solapadament de síntesi i emmagatzematge, secreció i proteòlisi del VWF.[167]

La VWD 2B està causada per mutacions *missense* en una regió discreta del domini A1 (exó 28).[160] Actualment, s'han reportat únicament 20 mutacions *missense* que afecten a 16 aminoàcids diferents, la majoria dels quals codifiquen per Arginina (Arg). Les més freqüents són p.Arg1341Gln, p.Arg1308Cys i p.Arg1306Trp.[133, 168] Les mutacions de tipus 2B provoquen canvis conformacionals del VWF afavorint la unió al receptor plaquetar GpIbα, el que es tradueix en un guany de funció que augmenta la unió a les plaquetes.[169]

La VWD 2M està causada per mutacions *missense* localitzades principalment al domini A1 i A3 el que condueix a una reducció de la unió del VWF a plaquetes o col·làgen.[17, 170] En concret, les mutacions que afecten a la unió amb GpIbα es troben al domini A1 (exó 28). D'altra banda, mutacions que afecten a la unió amb els col·làgens tipus I/III es localitzen als dominis A1 i A3 i les d'unió als col·làgens IV/VI únicament al A1, com la mutació p.Arg1399His.[171]

La 2N VWD presenta una herència autosòmica recessiva i els pacients que desenvolupen aquesta patologia són homozigots per una mutació *missense* de tipus 2N, o bé són heterozigots compostos degut a dues mutacions diferents de tipus 2N o una mutació 2N combinada amb un al·lel nul. Les mutacions de tipus 2N es localitzen als dominis D' (exó 17-20) i D3 (exó 24-25) i actuen disminuint la unió del VWF al FVIII que provoca un augment de la proteòlisi i aclariment del FVIII. Addicionalment, tot i que en menor freqüència, també s'han descrit mutacions a la regió C-terminal del propèptid

(D2), com la mutació p.Arg760Cys localitzada a l'exó 17 que impedeix l'escissió del propèptid del VWF i altera la unió del FVIII al VWF.[172]

5.2 Implicació funcional de les variants/mutacions identificades

Un repte fonamental en el diagnòstic molecular és determinar la rellevància biològica de les variacions, és a dir, demostrar que són responsables de la patologia i, per tant, es poden considerar mutacions. En aquest sentit, es disposa de diferents aproximacions que permeten ajudar a caracteritzar les mutacions com: la recerca de les descripcions prèvies de les mutacions en les bases de dades i literatura científica, estudis *in silico*, estudis del processament de l'mRNA *in vivo* i estudis d'expressió *in vitro*.

5.2.1 Definició de polimorfisme i mutació

En primer lloc, és molt important establir una bona definició que ens permeti diferenciar conceptualment què és una mutació (deletèria) i què un polimorfisme (no responsable de la patologia). No obstant, actualment no es disposa d'una definició vàlida en tots els casos per discriminar unes de les altres.

Una definició aproximada es basa en la freqüència de les variants a nivell poblacional, ja que, en la majoria de casos, es consideren polimorfismes aquells canvis en la seqüència de nucleòtids del DNA presents en més de l'1% de la població. Aquesta definició es recolza per la hipòtesi, gairebé sempre certa, que una major incidència en la població indica un efecte neutre o beneficiós. Ara bé, és remarcable que la presència d'una variable a la població sana no és excloent de patogenicitat. Un cas il·lustratiu són les variants p.Arg854Gln, p.Tyr1584Cys del *VWF* que tenen una freqüència >1% en població caucàsica però estan associades amb la VWD de tipus 2N i tipus 1, respectivament. Per tant, la freqüència poblacional és un paràmetre important per descartar la possibilitat que una variant concreta sigui deletèria, però no prova inequívocament que sigui benigna. En aquest sentit existeixen diferents bases de dades recullen la freqüència al·lèlica de les variants en diverses poblacions mundials (Caucàsica, Africana, Euroasiàtica, etc.): *Single Nucleotide Polymorphism*, dbSNP

(http:// www.ncbi.nlm.nih.gov/SNP); *Exome Variant Server*, EVS (http://evs.gs.washington.edu/EVS), *l'Exome Aggregation Consortium*, ExAC (http://exac.broadinstitute.org) i la base de dades del projecte 1.000 Genomes (http://www.1000genomes.org). Resulten destacables, pel gran volum de dades poblacionals que contenen: l'ExAC amb variants de 60.706 individus no relacionats (Setembre 2018) i; el projecte 1.000G que es va iniciar al 2007 amb la finalitat d'analitzar el genoma de 1.000 individus per estudiar la variabilitat genètica humana, i actualment ja n'han seqüenciat més de 3.000.

En base a aquestes premisses, l'*American College of Medical Genetics and Genomics*, l'*Association for Molecular Pathology* i el *College of American Pathologists* han establert uns criteris per determinar la patogenicitat de les variants detectades tenint en compte: 1) si la variant dóna lloc a una proteïna truncada (com ara mutacions *nonsense* o de canvi de pauta de lectura); 2) la co-segregació amb la malaltia en la família, 3) la descripció prèvia a literatura i/o bases de dades internacionals; 4) el resultat dels estudis *in silico* i; 5) la presència d'un segon al·lel mutat en el cas d'herència autosòmica recessiva. Aquesta avaluació permet classificar les variants en patogèniques, probablement patogèniques, de significat incert, probablement benignes o benignes.[173]

5.2.2 Descripció de la mutació en bases de dades

Un cop identificada la mutació en el pacient és imprescindible fer una cerca bibliogràfica de la variant per verificar si ha estat descrita prèviament en pacients amb patologia similar, fet que indicaria l'associació de la variant amb el desenvolupament de la patologia. En aquest sentit, existeixen bases de dades de mutacions locus específiques (LSMD) que recopilen mutacions d'una sola patologia. En el cas del *VWF* destaquen la del *VWF Scientific and Standardization Commite of the International Society of Thrombosis and Haemostasis* (SSC-ISTH) localitzada a la pàgina web de la *European Association for Haemophilia and Allied Disorders* (EAHAD) (https://grenada.lumc.nl/LOVD2/VWF/variants.php) amb un total de 708 entrades de mutacions (accés 16/03/2018) i l'Hemobase (www.hemobase.com) amb més de 120

mutacions, dissenyada al nostre laboratori.[174] Per altra banda, existeixen bases de dades que recopilen mutacions de totes les malalties hereditàries humanes, incloent la VWD, com *Human Gene Mutation Database* (www.hgmd.cf.ac.uk) i ClinVar (www.ncbi.nlm.nih.gov/clinvar).

5.2.3 Estudis *in silico*

Les mutacions de tipus *frameshift*, *nonsense* i grans reordenaments tenen un inqüestionable efecte patogènic. No obstant, en altres mutacions de tipus *missense*, sinònimes i intròniques el seu efecte patogènic no és tant evident. Per aquest motiu, s'han desenvolupat programes bioinformàtics que pretenen predir l'efecte patogènic d'aquestes mutacions de forma ràpida i econòmica, ja que la majoria d'aquests programes són de lliure accés .

Els programes que prediuen l'impacte de mutacins *missense* es basen principalment en la conservació filogenètica de la seqüència, dades d'estructura de la proteïna i en les propietats fisicoquímiques dels aminoàcids (natiu i mutant). Per exemple, els programes SIFT[175] o Mutation Assessor[176] utilitzen algoritmes basats en la conservació evolutiva del DNA, de manera que, un canvi en un aminoàcid molt conservat entre múltiples espècies és probable que esdevingui patogènic. Altres programaris com Polyphen-2[177] o Mutation Taster[178] es basen en la combinació de la informació de la seqüència i estructura de la proteïna per analitzar com la mutació afecta a les seves propietats fisicoquímiques (composició, polaritat i volum molecular). Per últim, el MutPred2 presenta un mètode basat en l'aprenentatge automàtic i un paquet de programari que integra dades genètiques i moleculars.[179]

Tanmateix, existeixen programaris per avaluar l'efecte de les potencials mutacions que afecten a l'*splicing* de l'mRNA. Les substitucions localitzades als dinucleòtids GT de l'extrem 5' i AG de l'extrem 3' dels introns es consideren clarament patogèniques. No obstant, mutacions *missense*, sinònimes o intròniques properes a la regió intró-exó poden afectar al processament de l'mRNA o crear nous llocs d'*splicing*. Per avaluar

aquest efecte el software més utilitzat es NNSplice,[180] NetGene2,[181] Human Splicing Finder[182] i MaxEntScan.[183]

Per avaluar l'efecte *in silico* d'una mutació és recomanable l'ús de mínim 3 anàlisi *in silico* basats en diferents algoritmes per donar robustesa a la predicció.[173] Tot i que, els resultats obtinguts a partir dels estudis *in silico* no han de ser considerats definitius a l'hora d'establir la patogenicitat d'una mutació, sinó que es requeriran estudis funcionals o d'altres evidencies per confirmar l'efecte deleteri.

5.2.4 Estudis del processament de l'mRNA del *VWF*

Amb la finalitat de determinar i/o confirmar l'efecte d'una PSSM en el processament de l'mRNA són necessaris la realització d'estudis transcripcionals. Atès que la biosíntesis *in vivo* del VWF està limitada a cèl·lules endotelials i megacariòcits,[184] gairebé tots els estudis utilitzen mRNA de plaquetes que, tot i ser anucleades, contenen petites quantitats d'mRNA del *VWF*.[185-188] No obstant, una altra font per obtenir mRNA són els leucòcits, on el *VWF* s'expressa de forma ectòpica.[189]

Les PSSM al *VWF* poden afectar el processament de l'mRNA*,* provocant la pèrdua d'un exó, la retenció d'introns o altres anormalitats de l'mRNA.[190, 191] Aquests efectes, en la majoria de casos, es tradueixen en la introducció de codons de parada prematurs (PTCs) en el marc de la pauta de lectura. Els PTCs són reconeguts i degradats de manera eficient per la cèl·lula en un procés nomenat *nonsense-mediated decay* (NMD), que afavoreix l'eliminació d'mRNAs aberrants. El mecanisme NMD presenta una eficiència diferencial segons el tipus cel·lular, sent més efectiu en el lloc d'expressió de la proteïna, en el cas del VWF, en plaquetes i cèl·lules endotelials.[192] Per aquest motiu, l'estudi en paral·lel de les mutacions mitjançant mRNA obtingut de leucòcits és recomanat ja que en ocasions permet identificar l'efecte real de les mutacions, que no s'observa a plaquetes degut a la degradació de l'al·lel per l'NMD.[191]

5.2.5 Estudis funcionals d'expressió del VWF

Els estudis funcionals mitjançant l'ús de models cel·lulars (*in vitro*) són , en ocasions, imprescindibles per determinar el mecanisme patogènic a nivell proteic de les mutacions identificades al *VWF*. Tradicionalment, aquest estudis s'han realitzat mitjançant transfeccions del construccions del *VWF* recombinant en línies cel·lulars heteròlogues. Les més utilitzades son les COS-7 (cèl·lules de ronyó de mona verda de Saba), les HEK293 (derivades de ronyó embrionari humà) i les AtT-20 (derivades de glàndula pituïtària del ratolí). La utilització d'aquestes línies cel·lulars ha permès la caracterització de mutacions responsables dels defectes d'unió del VWF a FVIII i a la GPIb. No obstant, aquest sistema presenta algunes limitacions, com per exemple, la dificultat d'establir un model heterozigot en les cèl·lules transfectades i els entrebancs observats en l'estudi de l'emmagatzematge i la secreció regulada del VWF, ja que aquestes línies cel·lulars no presenten orgànuls similars als WPB on s'emmagatzema el VWF (veure secció 3.2.2).[193]

Un model més apropiat per a l'estudi dels defectes en la proteïna del VWF són les cèl·lules endotelials ja que aquest és el lloc d'expressió natural del *VWF*. Els primers estudis d'aquest tipus es van realitzar amb cèl·lules endotelials de vena umbilical humana (HUVEC) aïllades de neonats amb VWD. [194-196] No obstant, degut a la dificultat per obtenir HUVECs, no és una opció factible en la majoria de casos. Per contra, en els últims anys s'han desenvolupat procediments per obtenir cèl·lules endotelials del pacient a partir d'una mostra de sang perifèrica, anomenades BOECs (de l'anglès *Blood Outgrowth Endothelial Cells*) que representa una valuosa alternativa per estudiar les mutacions del *VWF* en el seu entorn cel·lular natural (Figura 18).[197, 198]

Les BOECs són cèl·lules mare derivades de medul·la òssia que tenen la capacitat de diferenciar-se en cèl·lules endotelials madures.[199] Al 2013, dos estudis presentats per *Strake et al.* i *Wang et al.* van demostrar l'eficàcia d'aquestes cèl·lules en la investigació de 13 mutacions identificades en vuit pacients de tipus 1 VWD, tres de tipus 2M, un de tipus 2A i un de tipus 2N. Ambdós estudis ressalten els avantatges de les BOECs per

analitzar el comportament del *VWF* mutat en el tràfic intracel·lular, l'emmagatzematge i l'exocitosis.[200, 201]

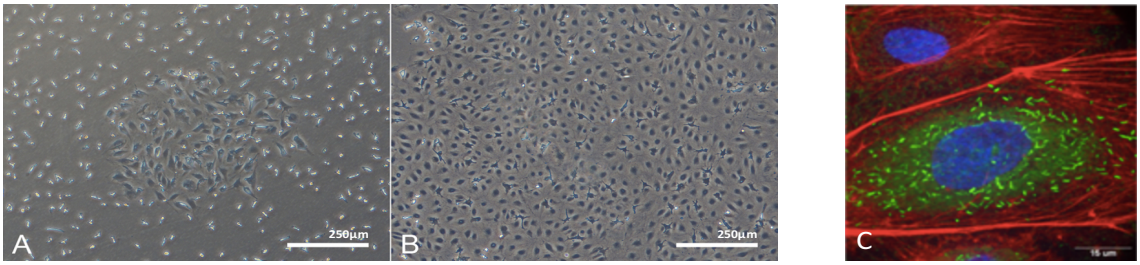


Figura 18. Visualització de BOECs d'un individu control. *A. Imatge d'una còlonia de BOECs. Les primeres colònies apareixen al cultiu entre els 7 i 14 dies. B. Imatge després del primer passe, les BOECs presenten una morfologia de llamborda pròpia de les cèl·lules endotelials. C. Imatge confocal d'immunofluorescencia de BOECs. En verd es mostra el VWF intracel·lular, en vermell l'actina i en blau l'àcid nucleic.*[197, 198]

6 Importància dels estudis de grans cohorts en la VWD

Els estudis poblacionals són la millor eina per a obtenir una visió àmplia de la complexitat de la VWD, i permeten aprofundir en el coneixement de les bases moleculars i del fenotip de la patologia així com fonamentar la relació genotip-fenotip per als tres tipus de VWD.

6.1 Cohorts de tipus 1

Els tres primers estudis multicèntrics de la VWD de tipus 1, presentats entre el 2006 i 2007, van permetre caracteritzar a nivell fenotípic i genotípic més de 300 *propositus* diagnosticats amb VWD de tipus 1 segons els criteris d'inclusió de la ISTH.[17] Aquests són el del Regne Unit, *United Kingdom Haemophilia Centre Doctors' Organization (UKHCDO),*[185] el canadenc en el qual van participar 13 centres locals[202] i l'europeu MCMDM-1VWD on van participar 14 centres de 9 països.[186] Tots ells han mostrat diverses característiques comunes sobre els detalls patogènics de la malaltia. En primer lloc, és destacable la identificació de mutacions candidates en únicament un 65% dels casos. En segon lloc, s'ha observat la identificació de més d'una mutació al *VWF* en un

15% dels pacients. Per últim, les variacions més freqüents identificades en aquestes cohorts són de tipus *missense* (70%), seguides de les mutacions d'*splicing* (9%), transcripció (8%), petites delecions (6%), *nonsense* (5%) i petites insercions o duplicacions (2%).[133]

Posteriorment, s'han presentat varis estudis de cohorts de la VWD de tipus 1 com l'estudi nacional suec on es va realitzar l'anàlisi genètic en 54 pacients diagnosticats al centre de referència de l'Hospital Universitari de Malmö.[203] Els resultats d'aquest estudi revelen que l'espectre de mutacions és similar al reportat en l'estudi europeu MCMDM-1VWD, tot i que s'han identificat 7 mutacions noves. Per altra banda, cal destacar el primer estudi realitzat en cohort pediàtrica de Canadà, compresa per 23 individus de 5 a 17 anys.[204] En 17 dels 23 pacients es van identificar un total de 14 mutacions, de les quals 4 no havien estat reportades prèviament. Per últim, l'estudi més extens en aquest tipus de VWD és el conegut com a programa Zimmerman (ZPMCB-VWD, *Zimmerman Program for the Molecular and Clinical Biology of VWD*), on han participat 31 centres d'Estats Units amb el reclutament de 482 pacients. Els resultats d'aquest estudi remarquen que la identificació de mutacions al *VWF* és més freqüents en individus que presenten nivells de VWF:Ag <30 IU/dL. [52]

6.2 Cohorts de tipus 2

La patologia genètica dels subtipus qualitatius de la VWD ha estat àmpliament estudiada per molts grups d'investigació. No obstant, en aquest àmbit s'han realitzat pocs estudis poblacionals amb un número significatiu de pacients, però se'n destaquen tres. En el primer estudi presentat al 1977 per la Xarxa francesa INSERM sobre anormalitats moleculars en VWD, es van reclutar 150 pacients.[205] Els resultats d'aquest estudi destaquen la identificació de mutacions candidates en tots els individus. A més, es van reportar 16 mutacions noves: quatre de tipus 2N als dominis D' i D3, set de tipus 2A als dominis A1 i A2, i cinc de tipus 2M. Per contra, no es van detectar mutacions noves causants de VWD de tipus 2B, el que destaca la poca heterogeneïtat al·lèlica en aquest subgrup. Al 2009, es va reportar l'estudi d'una cohort de 67 pacients de tipus 2B

VWD provinents de Milan, Vicenza i Utrech amb l'objectiu principal d'investigar la prevalença de la trombocitopènia, avaluar la gravetat dels sagnats i caracteritzar els pacients en termes genètics. Les mutacions més freqüents identificades en aquesta població van ser la p.Arg1306Trp (exó 15) i la p.Arg1308Cys (exó13).[206] Per últim, en un tercer es van reclutar un total de 56 pacients de l'Índia diagnosticats amb VWD de tipus 2. En aquesta cohort, la seqüenciació única del exó 28 va ser l'estratègia utilitzada per establir el diagnòstic molecular dels pacients i, degut a aquesta limitació metodològica, únicament van poder establir l'alteració molecular en 23 (41%) dels 56 pacients. En total, es van identificar 16 mutacions diferents (15 *missense* i 1 conversió gènica).[207] Els estudis de grans cohorts, juntament amb altres amb un número més reduït de pacients, com l'estudi canadenc de tipus 2M VWD realitzat en 16 pacients,[208] han permès millorar el coneixement de la patologia de la VWD de tipus 2 i aprofundir en la correlació genotip-fenotip.

6.3 Cohorts de tipus 3

Els primers quatre estudis genètics de cohorts de tipus 3 van permetre establir les bases moleculars d'aquesta patologia en 111 individus, aquests inclouen: l'estudi de 24 individus d'origen finlandès o suec; l'estudi de 27 individus d'origen indi o grec; l'estudi de 20 pacients del nord-oest d'Anglaterra i el més extens que inclou 40 pacients de diferents poblacions ètniques (italians, iranians i indis).[209-212] Posteriorment s'han publicat altres estudis com el canadenc on s'ha investigat a 34 *propositus*.[213] De tots ells, destaca el realitzat en 40 pacients on es van identificar mutacions en tots els al·lels, el que va conduir a una bona correlació genotip-fenotip en tots els casos.[212] En la resta de cohorts, la taxa de detecció de mutacions és del 85%, fet que s'explica per la possible presència de mutacions intròniques, en regions reguladores de l'expressió del *VWF* o de reordenaments no detectables amb l'estratègia de seqüenciació utilitzada.[209-211, 213]
En aquestes cohorts les mutacions més identificades, entre un 80 i 90%, són les que donen lloc a la generació d'al·lels nuls. Aquestes inclouen mutacions *nonsense* (31%), petites delecions i insercions (29%), PSSM (13%), grans delecions (6%) i conversions

gèniques (6%). No obstant, també s'han identificat mutacions *missense* (18%) que s'agrupen al propèptid i la regió carboxil-terminal del *VWF*.[209-213]

6.4 Registres nacionals europeus de la VWD

A Europa destaquen principalment tres grans registres nacionals on s'ha investigat la relació genotip i fenotip dels pacients. El primer es va presentar a Itàlia l'any 2006, conegut com a RENAWI, amb l'objectiu secundari de determinar els pacients i tipus de VWD que requerien tractament amb DDAVP i/o concentrats plasmàtics. En aquest estudi retrospectiu van participar 16 de 48 centres hospitalaris d'Itàlia i va resultar en el reclutament de 1.234 individus (451 famílies).[214] El segon, és el registre d'Alemanya on es van incloure 114 individus diagnosticats prèviament amb diferents tipus de VWD (78 famílies), el 80% provenien del centre d'hemofília de Bonn i el 20% restant d'altres regions del país.[215] Per últim, recentment, l'any 2016 s'ha publicat l'estudi prospectiu del registre francès on s'han analitzat 1.167 individus (670 famílies).[216] En els tres casos l'anàlisi genètic del *VWF* es va realitzar en els *propositus* de cada família mitjançant la seqüenciació de *Sanger*. Una bona correlació genotip-fenotip es va establir en un 84,6% dels *propositus* analitzats del registre alemany i en un 99,7% dels *propositus* del registre francès.[215, 216] En el registre RENAWI l'anàlisi exhaustiu del genotip i fenotip va permetre establir una bona caracterització dels pacients amb VWD en aquesta població i descriure la gestió d'aquesta malaltia a Itàlia.[214]

En conjunt, tots els estudis descrits han permès una millor comprensió de les bases moleculars dels diferents tipus de VWD, així com esbrinar els mecanismes fisiopatològics de les mutacions del *VWF*. Tanmateix, la correlació de genotip-fenotip analitzada en les cohorts han contribuït a millorar el diagnòstic i el tractament dels pacients amb VWD. Aquestes evidències avalen la importància dels estudis de grans cohorts per ampliar els coneixements d'aquesta patologia i millorar en última instància l'assistència als pacients.

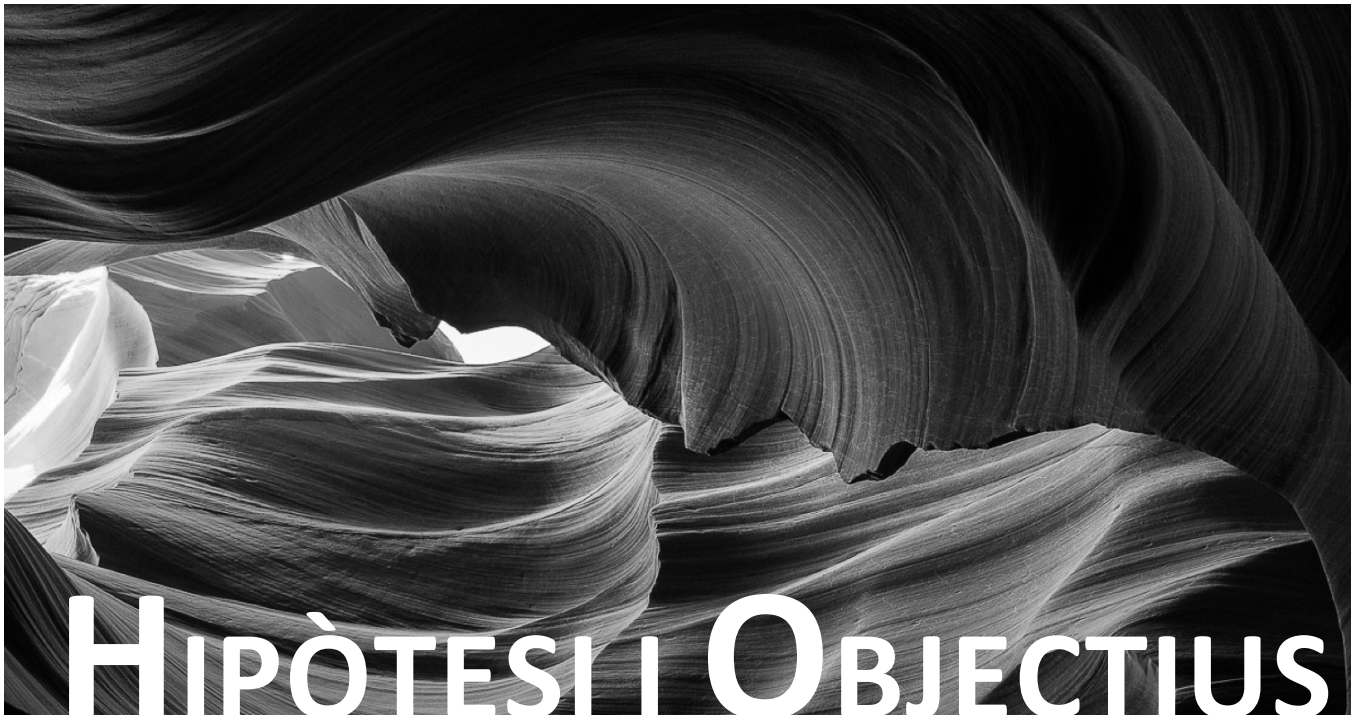
6.5 Presentació de la cohort espanyola i portuguesa de VWD

Motivat per l'èxit d'altres registres a nivell nacional o europeu de la malaltia de VWD com el "*Molecular and Clinical Markers in the Diagnosis and Management of VWD Type 1*'' (MCMDM-1 VWD),[186] al 2010 es va iniciar el projecte "Perfil clínic i molecular de pacients amb VWD: Registre Espanyol (PCM-EVW-ES)" amb l'objectiu de realitzar un registre nacional per a la VWD. Aquest projecte va suscitar l'interès de la comunitat d'hematòlegs espanyols degut a les dificultats que havien manifestat prèviament, a través d'un qüestionari, per establir un diagnòstic acurat i precís en un terç dels pacients a nivell nacional.[65] En aquest registre, pràcticament tots els grans hospitals espanyols (38), inclosos els cinc grans centres de referència d'hemofília, van participar en la selecció i reclutament dels pacients diagnosticats localment amb VWD, recollida de dades i diagnòstic clínic des del novembre del 2010 al 2013. En total es van reclutar 556 individus de 330 famílies.

Paral·lelament, en el mateix marc temporal, al país veí, Portugal, es va iniciar un projecte retrospectiu amb l'objectiu principal d'establir la correlació genotip-fenotip dels pacients diagnosticats amb VWD en aquesta regió. Per aquest estudi es van incloure 92 pacients de 60 famílies diagnosticats de VWD en 6 centres de 3 ciutats de Portugal (Lisboa, Faro i Coïmbra) entre els anys 2007 i 2014.

En ambdues cohorts de pacients es van centralitzar les mostres amb l'objectiu de realitzar l'anàlisi del genotip i fenotip de forma uniforme. En aquests sentit, els centres de referència seleccionats per a realitzar les proves de laboratori van ser l'Hospital Universitari i Politècnic La Fe de Valencia i el Complex Hospitalari Universitari A Coruña pel registre espanyol i l'Hospital i Centre Universitari de Coïmbra pel registre portuguès. D'altra banda, gràcies a l'experiència acumulada des del 1998 en el diagnòstic molecular de les coagulopaties congènites, el nostre laboratori va ser el responsable de seqüenciar i analitzar el *VWF* mitjançant procediments basat en NGS en ambdues cohorts.

En aquest context, el fonament d'aquest treball consisteix en estudiar la base genètica i la relació genotip-fenotip en el registre espanyol de la VWD, i participar en l'estudi genètic mitjançant la tecnologia NGS en la cohort portuguesa de pacients.



HIPÒTESI I OBJECTIUS

"El científico no es aquella persona que da las respuestas correctas, sino aquél quien hace las preguntas correctas"

Claude Lévi-Strauss

HIPÒTESI

Des de l'aparició de la NGS al 2005, la capacitat per amplificar, seqüenciar i analitzar el DNA ha millorat de forma exponencial. Tot i que inicialment aquesta tecnologia es va utilitzar principalment per a l'estudi de genomes, la reducció contínua dels costos i la simplificació dels procediments de preparació de mostres han promogut la seva aplicació, cada vegada en major mesura, al diagnòstic molecular de les malalties amb base genètica.

En base a aquestes premisses, la hipòtesi d'aquest treball postula que en el cas de la VWD, degut a la gran mida i complexitat del *VWF*, l'aplicació de la NGS a estudis genètics familiars o poblacionals resultarà especialment beneficiosa i ens permetrà investigar de forma exhaustiva les bases moleculars d'aquesta patologia. Tanmateix, pot esdevenir una eina indispensable pel diagnòstic, transformant l'escenari actual en el que l'anàlisi genètica és considerada una informació de suport, confirmatòria de la malaltia. Aquesta nova perspectiva ajudarà a establir més acuradament la relació de les mutacions amb la clínica hemorràgica i amb els resultats de laboratori observats, millorant la qualitat assistencial de la VWD.

OBJECTIUS

L'objectiu general d'aquesta tesi doctoral consisteix en aprofitar la tecnologia de NGS pel desenvolupament i optimització d'un protocol versàtil i econòmic per abordar l'estudi molecular de la VWD. Es pretén aplicar el nou abordatge a grans cohorts de pacients per aprofundir en la investigació dels mecanismes genètics de la VWD i determinar la implicació de les variants identificades al *VWF* en diferents aspectes del fenotip amb rellevància en el curs clínic de la malaltia. Els objectius concrets d'aquest treball es poden desglossar en:

- Dissenyar i optimitzar un protocol basat en la tecnologia de microfluids i NGS per a l'estudi molecular del *VWF* en grans cohorts de pacients diagnosticats de VWD.

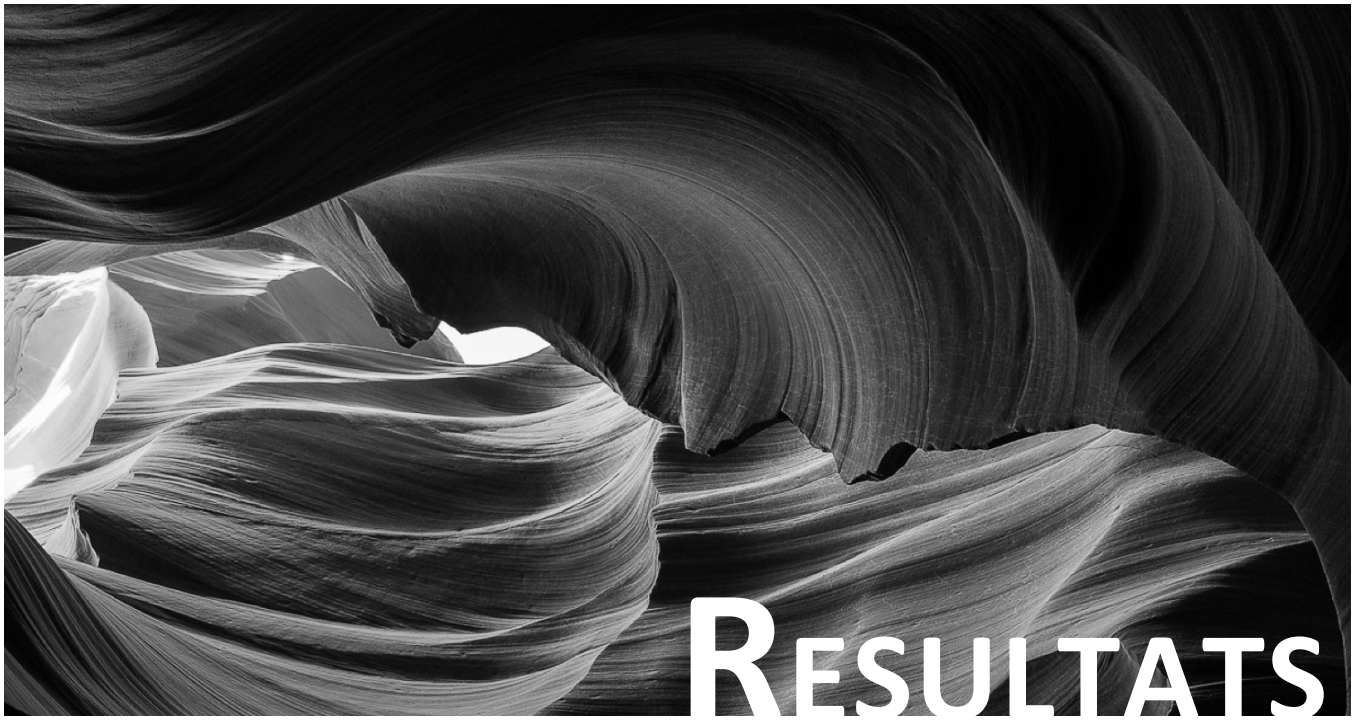
- Validar tècnicament el procediment dissenyat en mostres, prèviament estudiades per seqüenciació tradicional de *Sanger,* provinents de l'Hospital Universitari Vall d'Hebron i del Complex Hospitalari Universitari d'A Coruña.

- Aplicar el procediment a pacients reclutats de manera prospectiva en el registre espanyol de VWD (PCM-EVW-ES) com a eina de diagnòstic molecular i per investigar la correlació genotip-fenotip.

- Avaluar les potencials avantatges i inconvenients del procediment desenvolupat estudiant retrospectivament una cohort portuguesa de pacients amb VWD caracteritzats prèviament amb tècniques de seqüenciació tradicionals.

- Avaluar l'impacte funcional de les mutacions identificades al *VWF* mitjançant estudis *in silico* i estudis transcripcionals *in vivo* a partir de mostres dels pacients. Extrapolar els resultats obtinguts i valorar la contribució al fenotip.

- Estudiar els polimorfismes del *VWF* identificats en la cohort de pacients del registre PCM-EVW-ES per determinar el seu efecte sobre diferents trets diagnòstics de la VWD i esbrinar la seva potencial repercussió sobre la severitat de la patologia.



RESULTATS

"A la vida cal intentar moltes coses, moltes però moltes coses i moltes però moltes vegades; encara que algunes et surtin bé."

Miguel Àngel Cornejo

INFORME DELS DIRECTORS

Títol de la Tesi: Aplicació de les noves tecnologies de seqüenciació massiva al diagnòstic molecular de la malaltia de von Willebrand. Estudi de grans cohorts i anàlisi de la correlació genotip-fenotip.

Autora: Nina Borràs Agustí

Directors: Dr. Francisco Vidal Pérez i Dra. Irene Corrales Insa

La memòria de la Tesi Doctoral "Aplicació de les noves tecnologies de seqüenciació massiva al diagnòstic molecular de la malaltia de von Willebrand. Estudi de grans cohorts i anàlisi de la correlació genotip-fenotip" es presenta com un compendi de cinc articles científics, tres dels quals han estat publicats en revistes internacionals i dos estan sotmesos a publicació.

La contribució de la doctoranda Nina Borràs Agustí en cada una de les publicacions s'especifica a continuació, conjuntament amb el factor d'impacte i la categorització de cada revista.

ARTICLE 1

Batlle J, Pérez-Rodríguez A, Corrales I, López-Fernández MF, Rodríguez-Trillo Á, Lourés E, Cid AR, Bonanad S, Cabrera N, Moret A, Parra R, Mingot-Castellano ME, Balda I, Altisent C, Pérez-Montes R, Fisac RM, Iruín G, Herrero S, Soto I, de Rueda B, Jiménez-Yuste V, Alonso N, Vilariño D, Arija O, Campos R, Paloma MJ, Bermejo N, Toll T, Mateo J, Arribalzaga K, Marco P, Palomo Á, Sarmiento L, Iñigo B, Nieto M del M, Vidal R, Martínez MP, Aguinaco R, César JM, Ferreiro M, García-Frade J, Rodríguez-Huerta AM, Cuesta J, Rodríguez-González R, García-Candel F, Cornudella R, Aguilar C, **Borràs N**, Vidal

F. *Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): Proposal for a new diagnostic paradigm.* Thromb Haemost. 2016 Jan; 115(1):40-50. doi: 10.1160/TH15-04-0282.

Índex d'impacte: IF_{2016}: 5,627. 1r quartil (Q1) de la categoria *Hematology* (11/70) i Peripheral vascular disease (6/63).

Aportació de la doctoranda a l'article: La doctoranda, Nina Borràs, ha participat en l'anàlisi de les dades i la redacció del manuscrit publicat. Més concretament, el seu treball ha consistit en l'estudi de les regions pendents de cobrir i la recomprovació de les mutacions identificades en el *VWF* mitjançant la seqüenciació tradicional de *Sanger* per validar la fiabilitat i l'eficàcia de la tècnica basada en NGS. En aquells pacients on les mutacions identificades no explicaven el fenotip, ha realitzat MLPA amb l'objectiu de detectar grans duplicacions/delecions al *VWF*. Addicionalment, ha identificat mutacions al *F8* mitjançant l'aplicació d'un panell de gens en aquells pacients que es sospitava un diagnòstic d'hemofília A. En col·laboració amb els investigadors principals del Complex Hospitalari Universitari A Coruña, ha participat en la reclassificació i en l'establiment de la correlació genotip-fenotip de cada un dels pacients inclosos al registre.

ARTICLE 2

Borràs N, Batlle J, Pérez-Rodríguez A, López-Fernández MF, Rodríguez-Trillo Á, Lourés E, Cid AR, Bonanad S, Cabrera N, Moret A, Parra R, Mingot-Castellano ME, Balda I, Altisent C, Pérez-Montes R, Fisac RM, Iruín G, Herrero S, Soto I, de Rueda B, Jiménez-Yuste V, Alonso N, Vilariño D, Arija O, Campos R, Paloma MJ, Bermejo N, Berrueco R, Mateo J, Arribalzaga K, Marco P, Palomo Á, Sarmiento L, Iñigo B, Nieto MDM, Vidal R, Martínez MP, Aguinaco R, César JM, Ferreiro M, García-Frade J, Rodríguez-Huerta AM, Cuesta J, Rodríguez-González R, García-Candel F, Cornudella R, Aguilar C, Vidal F, Corrales I. *Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-*

ES): comprehensive genetic analysis by next-generation sequencing of 480 patients. Haematologica. 2017 Dec; 102(12):2005-2014. doi: 10.3324/haematol.2017.168765.

Índex d'impacte: IF_{2016}: 7,702. 1r quartil (Q1) de la categoria Hematology (4/70).

Aportació de la doctoranda a l'article: Aquest article exposa en detall els resultats dels estudis moleculars de la cohort presentada al primer article, de manera que l'aportació de la doctoranda, Nina Borràs, al treball de laboratori és l'exposat prèviament. A més, ha realitzat l'estudi *in silico* de totes les mutacions identificades al registre PCM-EVW-ES, així com l'anàlisi de les STRs en tots els pacients i familiars amb l'objectiu d'identificar un possible origen comú de les mutacions recurrents. També ha estat responsable de l'elaboració i revisió del manuscrit.

ARTICLE 3

Fidalgo T, Salvado R, Corrales I, Pinto SC, **Borràs N**, Oliveira A, Martinho P, Ferreira G, Almeida H, Oliveira C, Marques D, Gonçalves E, Diniz M, Antunes M, Tavares A, Caetano G, Kjöllerström P, Maia R, Sevivas TS, Vidal F, Ribeiro L. *Genotype-phenotype correlation in a cohort of Portuguese patients comprising the entire spectrum of VWD types: impact of NGS*. Thromb Haemost. 2016 Jul 4;116(1):17-31. doi: 10.1160/TH15-07-0604.

Índex d'impacte: IF_{2016}: 5,627. 1r quartil (Q1) de la categoria *Hematology* (11/70) i Peripheral vascular disease (6/63).

Aportació de la doctoranda a l'article: ha participat en la interpretació i validació dels estudis genètics. Més concretament, el seu treball ha consistit en l'anàlisi bioinformàtic dels resultats obtinguts de l'aplicació del procedimet basat en NGS per a l'estudi genètic del *VWF* en la cohort de 92 pacients de Portugal. D'altra banda, ha recomprovat per seqüenciació de *Sanger* les mutacions identificades en els pacients que presentaven resultats controvertits entre els obtinguts al Banc de Sang i Teixits mitjançant NGS i els obtinguts a l'Hospital i Centre Universitari de Coïmbra. Per últim, ha col·laborat en la

redacció i revisió del manuscrit final. Aquest article també forma part de la tesi doctoral de la primera signant, Dra. Teresa Fidalgo, defensada l'any 2017 a Coïmbra (Portugal).

ARTICLE 4

Nina Borràs, Gerard Orriols, Javier Batlle, Almudena Pérez-Rodríguez, Teresa Fidalgo, Patricia Martinho, María Fernanda López-Fernández, Ángela Rodríguez-Trillo, Esther Lourés, Rafael Parra, Carme Altisent, Francisco Vidal and Irene Corrales; on behalf of the members of the PCM-EVW-ES group. *Unraveling the effect of silent, intronic, and missense mutations on vwf splicing: contribution of NGS to mRNA study.* Haematologica (en revisió).

Índex d'impacte: IF_{2016}: 7,702. 1r quartil (Q1) de la categoria *Hematology* (4/70).

Aportació de la doctoranda a l'article: La doctoranda, Nina Borràs, ha participat en el disseny de la metodologia de l'estudi i ha realitzat tot el treball de laboratori de biologia molecular. Concretament, ha seleccionat els pacients amb mutacions candidates per a ser analitzades a nivell transcripcional, i ha realitzat l'extracció de l'RNA de leucòcits i plaquetes en cada un d'ells. Ha optimitzat el protocol prèviament utilitzat al laboratori per a l'amplificació de les regions d'interès del cDNA del *VWF* en cada pacient i ha dissenyat primers específics per a l'estudi de determinades mutacions. Els amplicons obtinguts els ha analitzat mitjançant seqüenciació de *Sanger* i NGS. Els resultats dels estudis *in vivo* els ha comparat amb la predicció obtinguda mitjançant quatre algoritmes *in silico* diferents fent una avaluació exhaustiva d'ambdues aproximacions. Per últim, ha estat la responsable de l'elaboració del manuscrit.

ARTICLE 5

Nina Borràs, Iris Garcia-Martínez, Javier Batlle, Almudena Pérez-Rodríguez, María Fernanda López-Fernández, Ángela Rodríguez-Trillo, Esther Lourés, Rafael Parra, Carme Altisent, Irene Corrales and Francisco Vidal; on behalf of the members of the PCM-EVW-ES group. *Characterization of the polymorphic variability of the VWF gene in the PCM-*

EVW-ES registry cohort and influence on phenotype. Journal of Thrombosis and Hemostasis (sotmès).

Índex d'impacte: IF_{2016}: 5,287. 1r quartil (Q1) de la categoria *Hematology* (13/70) i *Peripheral vascular disease* (8/63).

Aportació de la doctoranda a l'article: La doctoranda, Nina Borràs, ha recollit i tabulat totes les dades (clíniques y genètiques) emprades en la posterior anàlisi estadística. Ha participat en la realització de la estadística descriptiva de la cohort y la comparació amb altres registres moleculars amb dades del *VWF*. Ha col·laborat, aportant la seva visió biològica de la VWD als estudis estadístics basats en models de regressió lineal generalitzats mixtos per avaluar l'associació entre deu SNPs comuns i cinc mesures quantitatives relacionades amb VWD desenvolupats per la Dra. Iris Garcia-Martínez. Ha participat en la discussió i revisió final de les conclusions i en l'elaboració i revisió final del manuscrit.

Barcelona, setembre de 2018

Signatura dels directors

Dr. Francisco Vidal Pérez

Dra. Irene Corrales Insa

ARTICLE 1

Títol: *"*Perfil molecular i clínic de la malaltia de von Willebrand a Espanya (PCM-EVW-ES): proposta d'un nou paradigma de diagnòstic*".*

Resum

El diagnòstic de la malaltia de von Willebrand (VWD) continua sent difícil en una proporció important de pacients, representant a nivell espanyol un terç del total de casos. L'estudi multicèntric PCM-EVW-ES té com a objectiu investigar una cohort de 556 pacients (330 famílies) amb VWD de forma centralitzada. Amb aquest propòsit es van realitzar els test VWF:Ag, VWF:RCo, FVIII:C, VWF:CB i l'anàlisi de multímers en tots els pacients a nivell central. Tanmateix, l'anàlisi genètic de tota la regió codificant del gen del VWF (*VWF*) es va realitzar mitjançant un nou protocol basat amb la tecnologia de seqüenciació massiva (NGS). Els resultats obtinguts amb la reanàlisi del fenotip a nivell central van revelar una gran discrepància diagnòstica (42,3%) entre els valors dels centres locals envers als obtinguts a nivell central, fet que evidencia la dificultat diagnòstica d'aquesta patologia. L'anàlisi genètica en tots els pacients va resultar amb la identificació d'un total de 238 mutacions diferents al *VWF*, de les quals 154 no havien estat descrites prèviament a la base de dades internacional *Leiden Open Variation* (LOVD). Els resultats genètics es van comparar amb el fenotip per establir el diagnòstic final, que va permetre confirmar la patologia en 480 pacients. Una bona associació genotip-fenotip es va estimar en el 96,5% dels pacients. Cent setanta-quatre pacients presentaven dues o més mutacions. En 116 pacients es van identifica mutacions que prèviament s'havien associat amb un augment de l'aclariment del factor von Willebrand (VWF). L'estudi genètic va facilitar una classificació adequada en 63 de 70 pacients on el fenotip no va ser prou informatiu degut a la no disponibilitat del test RIPA, la discrepància dels resultats fenotípics centrals envers als locals i la difícil distinció entre la VWD de tipus 1 greu i tipus 3. En cinc pacients diagnosticats inicialment amb VWD però sense mutació identificada es va sospitar i/o confirmar una síndrome adquirida de

von Willebrand (AVWS). Aquestes dades recolzen el paper del diagnòstic genètic basat en NGS com a primera línia en el diagnòstic de la VWD.

Referència

Batlle J, Pérez-Rodríguez A, Corrales I, López-Fernández MF, Rodríguez-Trillo Á, Lourés E, Cid AR, Bonanad S, Cabrera N, Moret A, Parra R, Mingot-Castellano ME, Balda I, Altisent C, Pérez-Montes R, Fisac RM, Iruín G, Herrero S, Soto I, de Rueda B, Jiménez-Yuste V, Alonso N, Vilariño D, Arija O, Campos R, Paloma MJ, Bermejo N, Toll T, Mateo J, Arribalzaga K, Marco P, Palomo Á, Sarmiento L, Iñigo B, Nieto Mdel M, Vidal R, Martínez MP, Aguinaco R, César JM, Ferreiro M, García-Frade J, Rodríguez-Huerta AM, Cuesta J, Rodríguez-González R, García-Candel F, Cornudella R, Aguilar C, Borràs N, Vidal F.

Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): Proposal for a new diagnostic paradigm.

Thromb Haemost. 2016 Jan; 115(1):40-50. doi:10.1160/TH15-04-0282. PMID: 26245874

Molecular and clinical profile of von Willebrand disease in Spain (PCM–EVW–ES): Proposal for a new diagnostic paradigm

Javier Batlle[1]; Almudena Pérez-Rodríguez[1]; Irene Corrales[2]; María Fernanda López-Fernández[1]; Ángela Rodríguez-Trillo[1]; Esther Lourés[1]; Ana Rosa Cid[3]; Santiago Bonanad[3]; Noelia Cabrera[3]; Andrés Moret[3]; Rafael Parra[2,4]; María Eva Mingot-Castellano[5]; Ignacia Balda[6]; Carmen Altisent[4]; Rocío Pérez-Montes[7]; Rosa María Fisac[8]; Gemma Iruín[9]; Sonia Herrero[10]; Inmaculada Soto[11]; Beatriz de Rueda[12]; Víctor Jiménez-Yuste[13]; Nieves Alonso[14]; Dolores Vilariño[15]; Olga Arija[16]; Rosa Campos[17]; María José Paloma[18]; Nuria Bermejo[19]; Teresa Toll[20]; José Mateo[21]; Karmele Arribalzaga[22]; Pascual Marco[23]; Ángeles Palomo[5]; Lizheidy Sarmiento[24]; Belén Iñigo[25]; María del Mar Nieto[26]; Rosa Vidal[27]; María Paz Martínez[28]; Reyes Aguinaco[29]; Jesús María César†[30]; María Ferreiro[31]; Javier García-Frade[32]; Ana María Rodríguez-Huerta[33]; Jorge Cuesta[34]; Ramón Rodríguez-González[35]; Faustino García-Candel[36]; Rosa Cornudella[37]; Carlos Aguilar[38]; Nina Borràs[2]; Francisco Vidal[2]

[1]Complexo Hospitalario Universitario A Coruña, INIBIC, A Coruña, Spain; [2]Banc de Sang i Teixits, Barcelona, Spain; [3]Hospital Universitario y Politécnico La Fe, Valencia, Spain; [4]Hospital Universitari Vall d'Hebron, Barcelona, Spain; [5]Hospital Regional Universitario de Málaga, Málaga, Spain; [6]Hospital Universitario Dr. Negrín, Las Palmas de Gran Canaria, Spain; [7]Hospital Universitario Marqués de Valdecilla, Santander, Spain; [8]Salud Castilla y León, Segovia, Spain; [9]Hospital Universitario Cruces, Barakaldo, Spain; [10]Hospital Universitario de Guadalajara, Guadalajara, Spain; [11]Hospital Universitario Central de Asturias, Oviedo, Spain; [12]Hospital Universitario Miguel Servet, Zaragoza, Spain; [13]Hospital Universitario La Paz, Madrid, Spain; [14]Hospital Infanta Cristina, Badajoz, Spain; [15]Complexo Hospitalario Universitario Santiago de Compostela, Spain; [16]Hospital Universitario Lucus Augusti, Lugo, Spain; [17]Hospital Jerez de la Frontera, Cádiz, Spain; [18]Hospital Virgen del Camino, Pamplona, Spain; [19]Hospital San Pedro de Alcántara, Cáceres, Spain; [20]Hospital Sant Joan de Deu, Barcelona, Spain; [21]Hospital Sta Creu i St Pau, Barcelona, Spain; [22]Hospital Universitario Fundación Alcorcón, Madrid, Spain; [23]Hospital General de Alicante, Alicante, Spain; [24]Hospital Universitario 12 de Octubre, Madrid, Spain; [25]Hospital Clínico San Carlos, Madrid, Spain; [26]Complejo Hospitalario de Jaén, Jaén, Spain; [27]Fundación Jiménez Díaz, Madrid, Spain; [28]Hospital Nuestra Sra. de Sonsoles, Ávila, Spain; [29]Hospital Joan XXIII, Tarragona, Spain; [30]Hospital Ramón y Cajal, Madrid, Spain; [31]Hospital Montecelo, Pontevedra, Spain; [32]Hospital Río Hortega, Valladolid, Spain; [33]Hospital Gregorio Marañón, Madrid, Spain; [34]Hospital Virgen de la Salud, Toledo, Spain; [35]Hospital Severo Ochoa, Madrid, Spain; [36]Hospital Universitario Virgen Arrixaca, Murcia, Spain; [37]Hospital Lozano Blesa, Zaragoza, Spain;[38] Hospital Santa Bárbara, Soria, Spain

Summary

The diagnosis of von Willebrand disease (VWD) remains difficult in a significant proportion of patients. A Spanish multicentre study investigated a cohort of 556 patients from 330 families who were analysed centrally. VWD was confirmed in 480. Next generation sequencing (NGS) of the whole coding VWF was carried out in all recruited patients, compared with the phenotype, and a final diagnosis established. A total of 238 different VWF mutations were found, 154 were not included in the Leiden Open Variation Database (LOVD). Of the patients, 463 were found to have VWF mutation/s. A good phenotypic/genotypic association was estimated in 96.5 % of the patients. One hundred seventy-four patients had two or more mutations. Occasionally a predominant phenotype masked the presence of a second abnormality. One hundred sixteen patients presented with mutations that had previously been associated with increased von Willebrand factor (VWF) clearance. RIPA unavailability, central phenotypic results disagreement and difficult distinction between severe type 1 and type 3 VWD prevented a clear diagnosis in 70 patients. The NGS study facilitated an appropriate classification in 63 of them. The remaining seven patients presented with a VWF novel mutation pending further investigation. In five patients with a type 3 and two with a type 2A or 2B phenotype with no mutation, an acquired von Willebrand syndrome (AVWS) was suspected/confirmed. These data seem to support NGS as a first line efficient and faster paradigm in VWD diagnosis.

Keywords

VWD, VWF, NGS, phenotype, genotype

Correspondence to:
Francisco Javier Batlle Fonrodona, M.D.
Servicio de Hematología y Hemoterapia. INIBIC.
Complexo Hospitalario Universitario A Coruña
Edificio Hospital Materno Infantil
Carretera del Pasaje s/n
15006 – A Coruña, Spain
Tel.: +34 981 178000 Ext. 292113, Fax: +34 981 178392
E-mail: francisco.javier.batlle.fonrodona@sergas.es

† Jesús M. César died on August 25th, 2014.
Supplementary Material to this article is available online at www.thrombosis-online.com.

Financial support:
This study was supported by Baxter Healthcare Corporation for the PCM–EVW–ES support and Baxter BioScience Grant (H13–000845) for the genetic analysis. The Spanish Ministerio de Economía y Competitividad (MINECO)-Instituto de Salud Carlos III (ISCIII) (PI1201494 and RD12/0042/0053).

Received: April 6, 2015
Accepted after major revision: June 26, 2015
Epub ahead of print: August 6, 2015

http://dx.doi.org/10.1160/TH15-04-0282
Thromb Haemost 2016; 115: 40–50

influencing VWF multimer structure and interactions with platelet glycoprotein Ib (GPIbα), collagen and factor VIII (FVIII). An array of different mutation types producing either a null phenotype or abnormal VWF secretion is responsible for type 3 VWD. In contrast, the pathogenic mechanisms responsible for type 1 VWD remain only partially resolved.

VWD diagnosis requires a panel of assays and it is difficult in a significant proportion of patients because laboratory confirmation of VWD is a perilous journey (3–8). The measurement of VWF antigen (VWF:Ag) and VWF ristocetin cofactor activity (VWF:RCo) are key components in the diagnosis of VWD, although the introduction of direct platelet GPIbα-binding assays may become the functional assays of choice (1). Many individuals initially diagnosed with VWD do not have low or abnormal VWF on follow-up retesting. It is important to emphasise that the likelihood rate for diagnosis is usually improved if bleeding and family history and VWF levels are considered together (8–10). The assignment of some patients to the accepted categories may not be easy (2, 11, 12). The availability of different methodologies increases the problem. Also some tests, such as ristocetin-induced platelet agglutination (RIPA) and VWF multimeric analysis, are not generally available in all laboratories (13). Depending on the assay used some mutations may show normal values (14), or can only be detected by using types IV and VI collagen in patients who present with a clear mild bleeding history (15). It is also likely that other potential VWF defects causing VWD may remain undiagnosed (16).

The cost of laboratory diagnosis of VWD phenotype is becoming increasingly expensive. Currently, molecular genetic testing is recommended for types 2 and 3 VWD (8, 17). Newly developed genetic methods such as next generation sequencing (NGS) are becoming faster and progressively cheaper, allowing wider study of *VWF* mutations at a very reasonable cost (18).

A previous multicentre survey on VWD in Spain found diagnostic difficulties in at least one-third of patients (3). The present multicentre study investigated the characteristics of a large cohort of patients from Spain who were analysed centrally. NGS testing of the whole coding *VWF* was performed in all recruited patients and contrasted with the phenotypic data, obtaining a final diagnosis. One of the aims of this project was to investigate a possible role of the VWD genetic study in the initial VWD work-up.

Material and methods

Study design

The 'Molecular and clinical profile of von Willebrand disease in Spain (PCM–EVW–ES)' project covers a total of 38 nationwide centres. It was designed as reported previously (3), and began in November 2010. This study does fit neither in a prospective nor in retrospective categories. It is not just a conventional registry based in historical data, as long as it investigates (local and centrally) new collected samples, as already indicated in the text. We intend in the future to make a follow-up of the recruited patients in a prospective manner, and then it will be considered as prospective. The revised International Society on Thrombosis and Haemostasis (ISTH) classification (2) was used with some minor modifications (see Suppl. Material, available online at www.thrombosis-online.com).

PCM–EVW–ES online database Registry

Local and central data collected include, among many other parameters: socio-demographic information, type of bleeding episodes and a bleeding score (following a previously published questionnaire (19)), local and central phenotypic parameters, and central genotypic analysis.

Patients and controls

Patients of any age previously diagnosed locally with VWD between November 2010 and 2013 were recruited. All of them should fulfill one or more of the following inclusion criteria: 1) VWF:Ag, VWF:RCo and/or VWF:CB≤30 IU/dl (%), observed on two or more occasions; 2) detection of multimeric abnormalities; 3) in case of isolated FVIII deficiency it will be necessary to provide demonstration of a decreased VWF:FVIIIB; 4) presence of *VWF* mutations; 5) presence of RIPA at a low ristocetin concentration. The exclusion criteria were the presence of any data suggesting acquired Von Willebrand syndrome (AVWS). A 30 IU/dl VWF cut-off plasma level was used when a low VWF level was the only fulfilled criteria (see Suppl. Material, available online at www.thrombosis-online.com). As controls, 30 healthy subjects unrelated to any of the studied patients were also recruited for phenotypic analysis. Blood and plasma were collected and sent to three central laboratories.

This study was performed according to the guidelines of the Declaration of Helsinki. The study was approved by the local Research Ethics Committee, and all participants provided written informed consent.

Phenotypic assessment

The local phenotypic diagnosis was based on the following parameters: Platelet count and smear of peripheral blood (presence or absence of platelet aggregates), blood group, activated partial thromboplastin time (aPTT) , FVIII:C, VWF:Ag, VWF:RCo, and, optionally, RIPA to low concentration (0.4 mg/ml) and the closure times (CT) using the PFA-100® system (Platelet Function Analyser, Dade International, Miami, FL, USA) for collagen/ADP (C/ADP) and the collagen-epinephrine (C/Epi) cartridges. Centrally, FVIII:C, VWF:Ag, VWF:RCo and VWF collagen binding (VWF:CB) were measured in all recruited individuals' plasma samples. FVIII:C was assessed by a one-stage clotting assay using Actin FS as activator on a Behring Coagulation Timer (BCT) (Siemens Healthcare Diagnostics, Madrid, Spain).

The VWF:Ag was measured using the ELISA kit, DG–EIA VWF (Diagnostic Grifols, Barcelona, Spain). VWF:CB with collagen type III was determined by the ELISA kit DG–CBA VWF (Diagnostic Grifols), on a Triturus Immunoassay System. In some

samples VWF:CB was assayed by using type VI collagen (TECHNOZYM® VWF:CBA ELISA Collagen Type VI, Technoclone, Vienna, Austria).

The VWF:RCo test was performed on a BCT using the BC von Willebrand Reagent (Siemens Healthcare Diagnostics), and also by using conventional light transmission aggregometry.

The VWF capacity to bind exogenous FVIII (VWF:FVIIIB) was assessed using an in-house ELISA (20) in all patients who presented with a FVIII:C/VWF:Ag ratio < 1, and in those in whom a type 2N mutation was found. According to VWF:FVIIIB, the phenotype was classified as severe (< 0.3), moderate (0.3–0.78) and normal (0.79–1.22).

VWF multimeric distribution was analysed in all 556 patients and in the healthy controls using low (1%)- and high (2%)-resolution SDS-agarose gel electrophoresis, as described previously (21).

Lyophilised, reconstituted normal control plasma calibrated to the World Health Organisation standard was used as reference.

Genetic analysis

An automated method for *VWF* amplification using the Access Array™ platform (Fluidigm) was designed, developed and optimized. The previous NGS protocol for *VWF* analysis (18) was modified by adding the necessary specific adapters for NGS MiSeq Illumina sequencer. Detailed description about DNA extraction, primer design (Suppl. Table 1, available online at www.thrombosis-online.com) and PCR amplification are provided in the Suppl. Material (available online at www.thrombosis-online.com).

NGS sequencing of VWF and identification of genetic variants

The resulting amplification products from up to 192 patients were pooled and sequenced simultaneously in a MiSeq Desktop Sequencer (Illumina, San Diego, CA, USA). Barcoded sequences were de-multiplexed and analysed individually. The NGS pipeline output, paired sequence files (fastq format), was used as input for the analysis with CLC Genomic Workbench software (Qiagen, Hilden, Germany) and then with VariantStudio (Illumina). The optimal analysis parameters (coverage, minor allele counts, percent of variant allele, etc) were adjusted in order to obtain the optimal performance for mutation detection. This workflow allows the alignment of the resulting sequences against the human genome sequence (hg19) and concurrent *in silico* analysis, permitting the identification of potential pathogenic variants, discriminating pseudogene sequences and filtering known polymorphisms described to date in dbSNP and 1000 Genomes databases. Comparative analysis of sensibility and specificity with the traditional gold standard sequencing method showed values of nearly 99% and 98%, respectively. Furthermore, for the detection and assignment of the previously described mutations and polymorphisms, in addition to using the polymorphisms databases mentioned above, we reviewed the *VWF* international mutation database (22, 23) and the published literature to obtain previously described variants clinically associated with VWD.

Validation of the identified mutations by conventional Sanger sequencing

To confirm the mutations identified by NGS, the exact region mapping the mutation was amplified by PCR and sequenced by the dideoxynucleotide method, as described previously (24). The sequences obtained were aligned against the consensus wild-type *VWF* sequence using SeqScape (v2.7) software (Applied Biosystems, Foster City, CA, USA).

Multiplex ligation-dependent probe amplification (MLPA)

Samples from those patients in whom the detected mutations were unable to explain the clinical and/or laboratory phenotype were further analysed by the MLPA technique using the SALSA MLPA P011 and P012 VWF kits (version B2) (MRC–Holland, Amsterdam, The Netherlands) in order to detect deletions/duplications in *VWF*. This approach was applied for example in all type 3 patients in whom only one mutated allele was detected after sequencing and, likewise, in severe type 1 patients without mutation detected by NGS.

Fragment size analysis was performed using an ABI 3130 Genetic Analyser (Applied Biosystems). Data normalisation was carried out using four healthy controls and Coffalyser. Net software (MRC–Holland) was used for analysis using as input the fragment analysis files (.fsa) derived from the capillary Applied Biosystems system.

Genetic analysis of GP1BA

This was carried out by the Sanger method in patients with 2A or 2B multimeric pattern in whom no mutation was detected in *VWF* or when a new mutation was identified. The whole coding region of the *GP1BA* was amplified using primers as described previously (25).

Phenotype/genotype association

Once the genetic study was performed, the results were contrasted with the central phenotype and a final diagnosis was established. The phenotype/genotype association was evaluated according to the following criteria: 1) Excluded patients: none of the general inclusion criteria met, presence of a mutation not previously described with normal VWF and/or previously described mutation probably not causative (because associated with a normal phenotype); 2) Definite patients: clear phenotype correlated with a previously described mutation or with clear type 1 phenotype without mutation but with family history; 3) Candidate patients: clear phenotype associated with a variant not previously described. In this regard, a variant (as a genetic change) that was present in two or more patients with the same phenotype was considered to be

Table 1: Comparison between the local, central phenotypic diagnosis and final assignment according to the genetic and phenotypic association.

	Local n (%)	Central Phenotypic n (%)	Final n (%)
Type 1	285 (51.2 %)	126 (28.6 %)	133 (27.7 %)
Type 1H*	0	22 (5 %)	22 (4.6 %)**
Type 3	41 (7.4 %)	44 (10 %)	42 (8.8 %)
Type 3 carriers	2 (0.4 %)	5 (1.1 %)	30 (6.2 %)
Type 2	33 (6 %)	-	-
Type 2A	73 (13 %)	59 (13.3 %)	110 (23 %)
Type 2B	16 (2.9 %)	23 (5.2 %)	35 (7.3 %)
Type 2M	15 (2.7 %)	32 (7.2 %)	38 (7.9 %)
Type 2A/2M	24 (4.3 %)	38 (8.6 %)	34 (7 %)
Type 2N	26 (4.7 %)	9 (2 %)	9 (1.9 %)
Type 2N carriers	0	10 (2.3 %)	11 (2.3 %)
Compound heterozygous	1 (0.2 %)	4 (0.9 %)	-
Undiagnosed	40 (7.2 %)	70 (16.3 %)	16 (3.3 %)
TOTAL	**556**	**442**	**480**

** 1H: Type 1 historical (currently normal VWF properties but previously decreased).* ** One patient without mutation and twenty-one with mutation.

probably causative. Furthermore, patients categorised into type 2A/type 2B VWD with a mutation out of exon 28 were considered to be type 2A; 4) Undefined patients: VWD phenotype difficult to assign associated with a variant not previously described.

Results

Five hundred and fifty-six patients from 330 families historically diagnosed with VWD by their local centres were recruited into the project (►Table 1). After central phenotypic studies, VWD was confirmed in 442 patients; however, after genetic analysis 480 patients fulfilled some of the recruitment criteria.

Demographic data

Gender and blood group distribution in the cohort is shown in ►Figure 1.

Clinical data

A bleeding phenotype summary and bleeding score of the cohort are shown in ►Figure 2.

Local and central phenotypic diagnosis

Of 556 patients, 98 did not meet any inclusion criteria, three had an AVWS, eight had haemophilia A and five were carriers of a mutation in *F8* confirming VWD in 442 patients (►Table 1). Also, in 70 patients, it was necessary to wait for the genetic results, since with the phenotypic tests it was not possible to establish a clear diagnosis.

Only 484 out of 556 patients had a VWD type assigned locally. FVIII:C, VWF:Ag and VWF:RCo were performed locally in most patients. In addition, VWF:CB was assayed in 42 %, RIPA in 43 % and multimeric analysis in 19.2 %.

A local/central disagreement over diagnosis occurred in 205 out of 484 patients (42.3 %), and 33 additional patients were characterised as type 2 VWD without subtype specification. After the central phenotype study it was not possible to establish a clear diagnosis in 70 patients. Several reasons explain that finding. First, RIPA cannot be centralised and when it is not available at a local laboratory (as it occurred in 34 % of the participant centres) the assignment to type 2A, 2B or pseudo-VWD was not possible in 44 patients. Second, a central disagreement between the results of VWF:Ag and VWF:RCo versus multimeric analysis was present in 23 patients (5 %). Third, three patients could be either severe type 1 or type 3. The genetic analysis helped in the clarification in this regard of 63 patients. The remaining seven patients presented with a VWF novel mutation, which are pending further investigation (►Table 1).

VWF:FVIIIB was measured in 64 patients of whom 28 showed reduced binding (16 moderate and 12 severe). In the remaining 36 patients after performing the *F8* genetic study haemophilia A was confirmed in seven patients, five carriers and one pending of new study (Suppl. Table 2, available online at www.thrombosis-online.com). In 23 patients results were normal.

The results of the multimeric study are shown in ►Figure 1. ►Table 2 shows the median values and ranges of central laboratory data.

Genetic study

Genetic analysis was conducted for all patients included in the Registry (n = 556). For technical validation of the NGS methodology 172 patients previously studied by conventional sequencing techniques were analysed. The results confirmed that NGS allowed the detection of all previously detected mutations without prejudice to identify more mutations in some families. The only difficulty/discrepancy encountered in this regard was in the detection of a previously described large gene conversion (c.3835_4105conNG_001212.3:g.6566_6836; p.[(Val1279Ile;Gln1311Ter;Ile1343Val;Val1360Ala;Phe1369Ile)]). In this case, standard alignment parameters were inaccurate be-

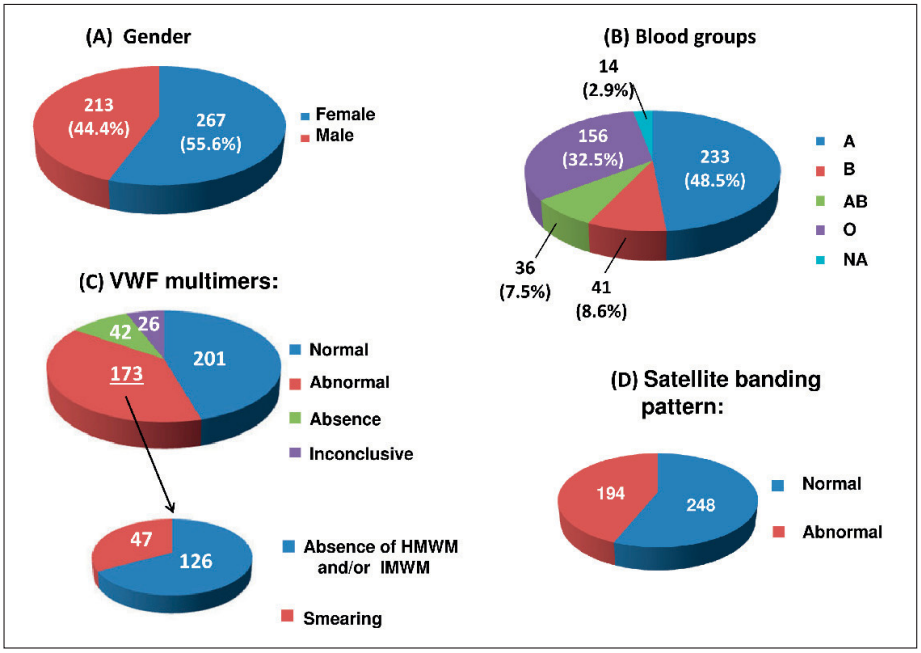
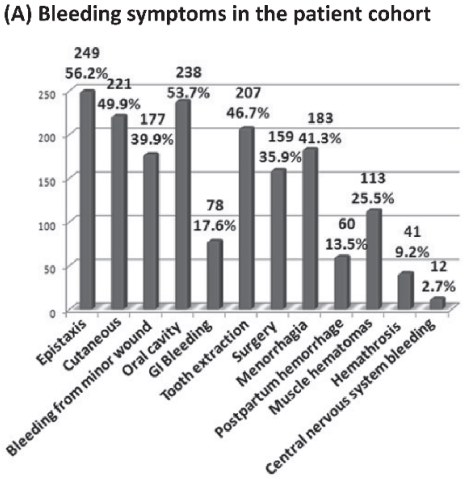


Figure 1: Demographic data of the cohort of patients. Gender (A) and blood group (B), and patients' distribution according to VWF multimer pattern in low (C) and high (D) resolution SDS-agarose gel electrophoresis. HMWM, high-molecular-weight multimers; IMWM, intermediate-molecular-weight multimers.

cause sequences preferentially aligned with the pseudogene and the mutation were not detected. In order to circumvent this problem, some parameters were changed and an additional alignment was performed to detect this kind of mutation.

After genetic analysis, 480 patients from 282 families were finally included in this Registry. A total of 463 patients presented with one or more mutations of *VWF*. Seven hundred twenty-three *VWF* changes due to 238 distinct mutations were detected: 84 had been described previously in the EAHAD Coagulation Factor Variant Databases VWF (the Leiden Open Variation Database, LOVD) (26) and there were 154 new candidate mutations. All types of point mutation, including some potential gene conversions, were identified along the entire gene (including flanking intronic re-



(B) Bleeding score (BS) after final diagnosis

VWD	BS (median)	Range
Type 3	12	(1, 26)
Type 2B	10	(0, 24)
Type 2A/2M	10	(0, 19)
Type 2A	10	(0, 23)
Type 2M	7	(-1, 21)
Type 2N	8	(3, 11)
Type 1	5	(0, 27)
Type 1H	3	(-2, 11)
Type 3 carrier	1	(-3, 10)
Type 2N carrier	0	(0,5)

Figure 2: Type of bleeding symptoms (in percentages) in the present cohort of patients (A) and median and range of bleeding score (BS) in the different types of VWD after final diagnosis (B).

Table 2: Median values and ranges of central laboratory data in patients according to the final diagnosis.

Central data									
VWD type	**FVIII:C (%)**	**VWF:Ag (%)**	**VWF:RCo (%)**	**VWF:CB (%)**	**VWF:C^{VI}B (%)**	**VWF:RCo/ VWF:Ag**	**VWF:CB/ VWF:Ag**	**VWF:C^{VI}B/ VWF:Ag**	**FVIII:C/ VWF:Ag**
1 (*n*=155)	41.30 (6.3–90)	21.59 (<4–60)	23.70 (<4–61)	19.73 (3.9–58)	-	1.10	0.91	-	1.91
2A (*n*=109)	36.42 (11.6–122)	28.83 (5.6–117)	8.10(<4–57)	12.84 (3.4–57)	-	0.28	0.44	-	1.26
2B (*n*=35)	50.63 (15–134)	43.60 (9–119)	27.79 (<4–103)	23.82 (5.7–71)	-	0.64	0.55	-	1.16
2A/2M (*n*=34)	32.88 (17–59)	22.99 (9.2–78)	9.86 (<4–31.8)	15.28 (5.8–34)	-	0.43	0.66	-	1.43
2M (*n*=38)	49.59 (11–213)	38.96 (5.4–272)	23.39 (<4–188)	28.53 (5.6–178)	-	0.60	0.73	-	1.27
2M* (*n*=8)	-	29.4 (6.8–84)	-	-	1.6 (0–6.14)	-	-	0.05	-
2N (*n*=9)	13.14 (4.5–30)	83.44 (32–274)	82.89 (36–167)	80.11 (26–267)	-	0.99	0.96	-	0.16
2N carrier (*n*=11)	81.36 (46–124)	69.73 (34–128)	89.91 (55–152)	74.54 (38–125)	-	1.28	1.07	-	1.17
3 (*n*=42)	4.98 (1.3–13)	1.18 (0.5–5.1)	0.72 (0.5–4.2)	0.85 (0.5–4.3)	-	0.61	0.72	-	4.22
Normal range	**60–140**	**47–19**	**50–170**	**60–130**	**(60–130)**	**>0.7**	**>0.7**	**>0.8**	**>1**

2M* (eight patients). VWF:^{CVIB}: Collagen type VI binding assay.

gions and promoter). For details, see Suppl. Results and Suppl. Figure 1 (both available online at www.thrombosis-online.com).

Definite or putative mutations were not detected in 17 patients who had quantitative defects exclusively: 16 type 1 and one type 1H.

MLPA analysis was performed in seven patients without mutation (four type 3, one severe type 1, one 2A/2B and one 2M) and in three patients with a single putative mutation (one severe type 1 and two type 3 patients). Alterations in the MLPA patterns, indicative of a large deletion, were detected in one type 3 and in one type 2A patient.

Mutations in type 1 patients

A total of 212 changes/mutations were found in 155 type 1 patients (including type 1H). Fifty-five patients had synonymous, intronic or upstream (promoter) variants. Three unrelated patients shared an identical synonymous mutation (c.3390C>T), adjacent to the 5´ end of exon 26 which could create a new potential donor splicing site.

When we analysed patients with type 1H, we detected 19 mutations, and only one patient had no mutation. In type 1H the percentage of intronic, synonymous and promoter changes increased to 37% of the total (26% when type 1 is considered altogether). Only in one patient was an intronic change found as a potentially exclusive responsible mutation.

Mutations in type 2 patients

Mutations found in type 2 patients were largely (80%) of the missense type and 219 out of 226 patients studied had at least one missense mutation, with 186 of the mutations being found in exon 28 (Suppl. Figure 2, available online at www.thrombosis-online.com). Among the seven remaining patients, one had a large in-frame heterozygous deletion comprising exons 11 to 32 (p.Gly386_Ser1873del), three patients had a small in-frame deletion (p.Lys1408del), and two patients had a small in-frame insertion (p.Asp1560dup) in exon 28. Furthermore, in the remaining patient we found one synonymous and two intronic variants.

Mutations in type 3 patients

Ninety-five changes/mutations were found in 42 type 3 patients included in the project. These were mainly missense (24) , nonsense (20) and splice site (12) mutations and were putatively responsible for the severe quantitative phenotype in our cohort (Suppl. Figure 3, available online at www.thrombosis-online.com). Among missense mutations, only p.Asn2066Ser was encountered several times (n = 6) but was always associated in *cis* with the frameshift mutation p.Tyr126ThrfsTer49; however, the real contribution of such a mutation to the phenotype has not yet been confirmed. In four patients we found the nonsense mutation p.Gln1311Ter always in the homozygous state. Eleven splicing mutations discovered in type 3 patients affected one of the two strictly consensus nucleotides of acceptor (AG) or donor (GT) splice site and one mutation affected nucleotide +3 of intron 50 (c.8155+3G>C). Nevertheless, mutations in type 3 patients were not always null or splicing mutations: two patients were homozygous for a novel missense mutation (p.Cys370Tyr) and for a previously described (27) controversial variation (p.Pro1162Leu).

Mutations in excluded patients

Finally, since we sequenced all individuals initially enrolled in the study we also have the variations of those who were later excluded from the PCM–EVW–ES project. We found changes in 24 individ-

mild haemophilia A in eight patients and a haemophilia A carrier status in five patients; 2) fifty-six patients did not meet any criteria for inclusion; 3) in seven cases without *VWF* and *GP1BA* mutations, an AVWS was suspected, which was confirmed in three patients. This highlights the fact that some patients with AVWS may initially be overlooked, and it is clinically important to confirm a previously unsuspected syndrome.

As far as type 1 VWD is concerned, seventeen out of 155 patients (10.9%) presented with no mutation and the possible explanations are: 1) a potential defect in a modifier gene (33, 34); 2) an unsuspected AVWS (35, 36); 3) presence of an undetected *VWF* mutation/s.

The yield of mutations observed in patients with type 1 in this study is significantly higher than that published previously (27–31). The cut/off VWF level of 30 IU/dl used as an inclusion criterion probably increases the likelihood of finding mutations.

Twenty-two cases were considered to be type 1H. This is important since type 1 patients evolve towards the normal range of VWF as they age, but they still present with clinical symptoms, and should not be considered as fully normal with regard to haemostasis (37).

With regard to type 3 VWD, seventy-two patients presented with some type 3 mutations and 42 of them showed a type 3 phenotype (25 homozygous and 17 heterozygous). Three additional double heterozygous patients with the mutations c.5170+10C>T plus c.7730–1G>C in *cis* showed a type 1 or a normal phenotype. Type 3 is regarded as a classic recessive trait, but a detailed inspection of type 3 families has shown that different manifestations of the haemostatic defect and haemorrhagic phenotype are common in individuals who are heterozygous carriers of certain type 3 mutations (38). In fact, approximately 50% of type 3 reanalysed families exhibit a typical recessive inheritance pattern while the remaining cases appear to show characteristics of co-dominant inheritance, with one or both parents showing clinical and laboratory evidence of type 1 VWD (38).

In our cohort only one mutation was significantly recurrent in this group: the p.Gln1311Ter mutation (derived from a potential gene conversion) which was found in eight patients (including type 3 carriers) from six unrelated families. This mutation had been identified previously in Spanish Romani patients with a probable shared ancestral origin and expected high frequency in our population (39).

There is a similar variability for the in-frame mutations detected in six unrelated families that seem to respond to singular and independent molecular events, as evidenced by the fact that three of them were present in one family and only one (p.Pro1079_Tyr1080insLeuGlnValAspProGluPro) was shared by two unrelated families.

With respect to types 2A, 2B and 2M VWD, RIPA is unavailable in a significant number of centres, making it difficult to distinguish between types 2A, 2B and pseudo-VWD. In 44 patients in whom RIPA was not assayed, the NGS analysis confirmed 25 mutations of type 2A, 10 type 2B, two patients with no mutation (with an AVWS confirmed in one of them) and seven patients with *VWF* novel mutations, pending further investigation. Eleven out of 12 patients with the type 2B mutation (p.Arg1308Cys) showed enhanced RIPA, while one patient showed a positive RIPA only at 0.8 mg/ml or higher concentrations. Analysis of *GP1BA* did not detect any mutation, and to our knowledge no patient with pseudo-VWD has been found in Spain.

Collagen binding heterozygous mutations were present in 11 patients, related to collagen types I and III (two patients), type IV (one patient) and type VI (eight patients). A VWF:C^{VI}B analysis in the latter patients confirmed a severe decrease in this activity (►Table 2), in contrast to a previous observation showing a severe decrease only in homozygous patients (8). Three patients from two different families showed the novel heterozygous mutations (p.Ala2498Asp, p.Cys2491Arg, respectively) in exon 44 responsible for the VWF-GP αIIbβ3 binding (29). These mutations are pending demonstration of their causative role and clinical impact. This function as well as those evaluated with types IV and VI collagens are not included in routine laboratory assessment.

The mutation p.Gly1415Asp, previously described as type 1 (28), was found in five patients from three unrelated families showing a phenotype compatible with type 2M.

In several patients with a phenotype that was difficult to classify, the genetic study facilitated an appropriate classification. Thus, 23 patients showing a non-clear multimeric pattern presented with a classical type 2A mutation. Likewise, several patients with the same mutation (sometimes from the same family), showed a variation in the degree of multimeric abnormality.

As far as Type 2N VWD is concerned, 28 patients with a 2N mutation were observed: 12 with severe phenotype (nine homozygous and three heterozygous but they presented with a severe phenotype due to the association with a second deleterious mutation). The remaining 16 patients were heterozygous and they presented with a moderate phenotype. Curiously, in three patients from two different families a single and new mutation (p.Arg762Ser) was associated with a compound phenotype (2A/2M) and moderate type 2N. A compound phenotype 2N and 2A due to a single mutation (p.Asn879Asp) was described previously (40). In a patient type 3 a type 2N carrier status was detected by NGS.

Interestingly, in those patients in whom a heterozygous type 2N was not suspected initially and a type 2N mutation was found, VWF:FVIIIB was abnormal. Conversely, when FVIII:C/VWF:Ag ratio was low and VWF:FVIIIB was normal no type 2N mutation was detected.

A single patient with 2A and 2N combined phenotypes was compound heterozygous for the p.Cys2773Ser and p.His817Gln mutations. Some authors suggested that p.His817Gln could be a polymorphism (27). However, others have considered a plausible type 2N variant as it lies within the FVIII binding region resulting in a significant decrease in FVIII binding capacity (41).

On the other hand, 116 patients presented with mutations that have previously been associated with an increased VWF clearance (►Table 4), which may preclude the use of desmopressin (DDAVP) (42).

In addition, 174 patients presented with two or more mutations in which a predominant phenotype could mask the presence of a

second defect. This has important implications for genetic counseling (e.g. a type 2N or 3 carrier) (▶Figure 4).

Expression of selected novel *VWF* mutations found will be carried out in the near future to confirm their causative role.

The role of NGS in VWD diagnosis

The first use of NGS was closely linked to the sequencing of the complete genome. Interest in its use in diagnostic applications has increased in recent years (43), especially in the analysis of diseases caused by large and highly polymorphic genes such as VWD (44). NGS seems to be essential when conducting population-based studies such as the PCM–EVW–ES (18).

Obviously, having the sequence of *VWF* from patients and families is advantageous to establish a more reliable diagnosis and also to detect various degrees of the disease that would otherwise be missed. The VWD phenotypical study presents several drawbacks such as low sensitivity of VWF:RCo and discrepancies of results obtained with different VWF activity assays (14, 27, 45–47). In theory, as yet undetected mutations other than collagen IV and VI may be associated with VWD in patients with a mild bleeding phenotype in whom current routine laboratory investigation is normal.

It has been recently reported that several frequent and rare variants described as polymorphisms, such as p.Thr789Ala, can have an effect on the VWF and/or FVIII levels (41). Moreover, there is an evident ethnic variability on VWF levels, probably due to the highly polymorphic VWF coding sequence. For instance, VWF:Ag and FVIII:C are higher in African Americans (AAs) than in European Americans (EAs). Likewise, the ratio of VWF activity (ristocetin cofactor activity) to VWF:Ag is lower among AAs than EAs. Consequently, AAs are diagnosed more frequently with the qualitative type 2M variant of VWD. This emphasises the importance of studying ethnically diverse populations to expand our knowledge of normal genetic variations and to distinguish pathogenic variants from non-deleterious polymorphisms (41). Thus, it seems indispensable further investigation to unravel the real effect of missense changes identified along the *VWF*. The rapid cost reduction achieved through the NGS could strongly support this new scenario: the combination of massive sequencing for mutation analysis with phenotypic and family data in a well defined population, as presented here, provides a dynamic image to investigate the exact contribution of the so called polymorphisms which may also subtly modulate the phenotype. It will lead to a better understanding of the mechanisms involved in the pathophysiology of VWD.

Furthermore, several physiological factors influence plasma VWF levels, requiring repeated quantification. Different studies have identified modifier genes involved in regulating plasma levels of VWF (34, 48–52) that may explain decreased VWF levels in some patients with type 1. The potential involvement of such genes outside of the *VWF* is unknown at present. They could account for a significant percentage of cases of VWD for which no mutation has been identified in the coding region of *VWF*. Thus, the success of mutation detection after *VWF* analysis in type 1 patients can never be expected to reach 100%.

Table 4: Number of patients who present mutations associated with increased clearance of VWF.

VWD type	n	Mutations (*n*)
1	40	p.Arg1205His (35); p.Va1760Ile (3); p.Ser2179Phe (2)
2A	14	p.Cys1149Arg (11); p.Leu1307Pro (3)
2B	28	p.Arg1308Cys (12); p.Arg1306Trp (7); p.Val1316Met; p.Pro1266Leu (2); p.Arg1306Gln (2)
2M	5	p.Gly1415Asp (5)
2A/2M	29	p.Arg1374Cys (22); p.Arg1315Cys (7)
Total	**116**	

Some investigators believe that, unlike the utility of genetic testing in type 2 and 3 VWD, such testing is much less clear in patients with type 1 (1, 5, 8, 17, 28). In contrast, the data provided here probably support another more efficient and faster strategy in VWD diagnosis by the inclusion of the NGS in the study as a first-line diagnostic method. The reasonable cost of NGS ($70/sample, even cheaper than most of the phenotypic tests) strongly supports this new scenario. It may also provide useful clinical information (such as the prediction of an increased VWF clearance), and it is not affected by all the previously mentioned drawbacks.

In conclusion, we wish to propose an algorithm for the initial study of VWD that includes family history, bleeding score, VWF:Ag, VWF:RCo/VWF:GPIbα assay, platelet count and peripheral blood smear, and *VWF* study by NGS, using a cut-off for VWF:Ag of ≤30 IU/dl in type 1 VWD, considering all of them as essential information to predict and manage the severity of the clinical bleeding phenotype. In a second step, some other mutation-oriented tests or investigations could be carried out for confirmation. Thus, this does not intend at all to transmit the message of molecular analysis as a surrogate approach of a phenotypic study. We are now developing an algorithm platform as a software

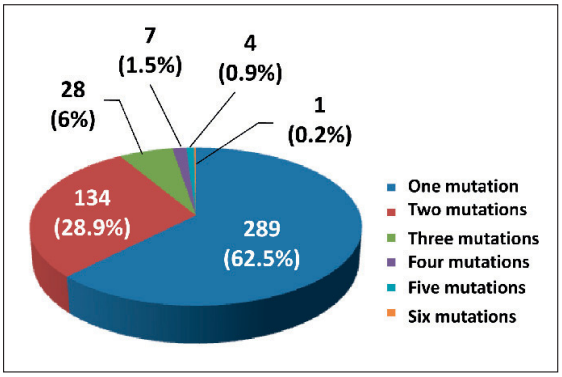


Figure 4: Distribution of the patients according to the number of mutations.

What is known about this topic?

- VWD diagnosis is difficult in a significant proportion of patients.
- The cost of laboratory diagnosis of VWD phenotype is becoming increasingly expensive.
- NGS is becoming faster and progressively cheaper, allowing wider study of *VWF* mutations at a very reasonable cost.

What does this paper add?

- A 42.3 % local/central phenotypic diagnosis discrepancy was observed indicating the practical difficulties/drawbacks of VWD diagnosis.
- A 96.5 % phenotypic/NGS genotypic association was observed, suggesting the possibility of incorporating the NGS in the initial VWD work-up.
- Comparative analysis of sensibility and specificity with traditional gold standard sequencing method showed values of nearly 99 % and 98 %, respectively.

based on these parameters that will help in the diagnostic process in the clinical setting. It includes a *VWF* mutations database that will be updated regularly, taking into account the *VWF* sequence variants that are known not to influence VWF or cause a clinical disease (22, 23, 27). This algorithm platform will be validated by using data from this project.

Acknowledgements

We are indebted to Baxter Healthcare Corporation for the PCM–EVW–ES support as well as the Baxter BioScience Grant (H13–000845) for the genetic analysis. This work was supported by the Spanish Ministerio de Economía y Competitividad (MINECO)-Instituto de Salud Carlos III (ISCIII) (PI1201494 and RD12/0042/0053). We thank Technoclone GmbH, Austria, Labclinics, SA, Spain the donation of a VWF:CB (type VI collagen) kit, and Dr. L. Rivero and Dr. T. Fidalgo, University Hospital, Coimbra, Portugal) for providing us with a sample from a patient with pseudo-VWD (D235Y mutation) as a positive control. Finally, we also deeply thank the kind participation of the recruited patients and to Ros Kenn for the medical editing of this manuscript.

Author contributions

Contribution: J. B. initiated and coordinated the study; all the authors participated in study design, data collection and performance of laboratory analyses; J. B., A. P. R., F. V., I. C., A. R. C., A. R. T., E. L., N. C. and A. M. analysed and interpreted results; J. B., A. P. R., F. V., I. C. and A. R. C. wrote the paper; and all the authors checked the final version of the manuscript.

Conflicts of interest

None declared.

References

1. Lillicrap D. von Willebrand disease: advances in pathogenetic understanding, diagnosis, and therapy. Blood 2013; 122: 3735–3740.
2. Sadler JE, Budde U, Eikenboom JC, et al. Update on the pathophysiology and classification of von Willebrand disease: a report of the Subcommittee on von Willebrand Factor. J Thromb Haemost 2006; 4: 2103–2114.
3. Batlle J, Perez-Rodriguez A, Pinto JC, et al. Diagnosis and management of von Willebrand disease in Spain. Semin Thromb Haemost 2011; 37: 503–510.
4. Costa-Pinto J, Pérez-Rodríguez A, del C Goméz-del-Castillo M, et al. Diagnosis of inherited von Willebrand disease: comparison of two methodologies and analysis of the discrepancies. Haemophilia 2014; 20: 559–567.
5. Castaman G, Goodeve A, Eikenboom J, et al. Principles of care for the diagnosis and treatment of von Willebrand disease. Haematologica 2013; 98: 667–674.
6. Favaloro EJ, Bodó I, Israels SJ, et al. von Willebrand disease and platelet disorders. Haemophilia 2014; 20 (Suppl 4): 59–64.
7. Federici A. Therapeutic Progress in von Willebrand Disease. Hematology Am Soc Hematol Educ Program; 2014: 524–530.
8. Flood VH. Perils, problems, and progress in laboratory diagnosis of von Willebrand disease. Semin Thromb Haemost 2014; 40: 41–48.
9. Tosetto A, Castaman G, Rodeghiero F. Evidence-based diagnosis of type 1 von Willebrand disease: a Bayes theorem approach. Blood 2008; 111: 3998–4003.
10. Flood VH, Gill JC, Friedman KD, et al. Von Willebrand disease in the United States: a perspective from Wisconsin. Semin Thromb Haemost 2011; 37: 528–534.
11. Batlle J, Pérez-Rodríguez A, Franqueira MD, et al. Type 2M von Willebrand disease: a variant of type 2A? J Thromb Haemost 2008; 6: 388–390.
12. Penas N, Pérez-Rodríguez A, Torea JH, et al. von Willebrand disease R1374C: type 2A or 2M? A challenge to the revised classification. High frequency in the northwest of Spain (Galicia). Am J Hematol 2005; 80: 188–196.
13. Hamilton A, Ozelo M, Leggo J, et al. Frequency of platelet type versus type 2B von Willebrand disease. An international registry-based study. Thromb Haemost 2011; 105: 501–508.
14. Ahmad S, Ptashkin B, DiGiovanni C, et al. False normal von Willebrand factor activity by monoclonal antibody-based ELISA in a patient with type 2A(IID) von Willebrand disease. Thromb Haemost 2011; 106: 1224–1225.
15. Gill JC, Haberichter SL, Bellissimo DB, et al. Binding Of VWF To Type IV Collagen: An Additional Collagen Binding Mechanism Beyond Types I, III, and VI Collagen? Blood 2013; 122: 333.
16. Flood VH, Gill JC, Christopherson PA, et al. Critical von Willebrand factor A1 domain residues influence type VI collagen binding. J Thromb Haemost 2012; 10: 1417–1424.
17. Flood VH. New Insights into Genotype and Phenotype of von Willebrand Disease. 2014; pp. 531–535.
18. Corrales I, Catarino S, Ayats J, et al. High-throughput molecular diagnosis of von Willebrand disease by next generation sequencing methods. Haematologica 2012; 97: 1003–1007.
19. Tosetto A, Rodeghiero F, Castaman G, et al. A quantitative analysis of bleeding symptoms in type 1 von Willebrand disease: results from a multicentre European study (MCMDM-1 VWD). J Thromb Haemost 2006; 4: 766–773.
20. Casonato A, Pontara E, Zerbinati P, et al. The evaluation of factor VIII binding activity of von Willebrand factor by means of an ELISA method: significance and practical implications. Am J Clin Pathol 1998; 109: 347–352.
21. Batlle J, López Fernández MF. Laboratory Assays for von Willebrand Factor. In Coagulation and Bleeding Disorders: The role of Factor VIII and von Willebrand Factor. Ed Marcel Dekker Inc New York 1989: pp. 325–342.
22. European Association for Haemophilia and Allied Disorders (EAHAD). Coagulation Factor Variant Databases. Available at: https://www.grenada.lumc.nl/LOVD2/VWF/home.php?select_db=VWF. Accessed November 25, 2014.
23. International Society on Thrombosis and Haemostasis-Scientific and Standardisation Committee of von Willebrand Factor (ISTH-SSC). VWF Online Database (VWFdb). Available at: http://www.vwf.group.shef.ac.uk/vwd.html. Accessed November 14, 2014.
24. Corrales I, Ramírez L, Altisent C, et al. Rapid molecular diagnosis of von Willebrand disease by direct sequencing. Detection of 12 novel putative mutations in VWF gene. Thromb Haemost 2009; 101: 570–576.
25. Othman M, Notley C, Lavender FL, et al. Identification and functional characterisation of a novel 27-bp deletion in the macroglycopeptide-coding region of

the GPIBA gene resulting in platelet-type von Willebrand disease. Blood 2005; 105: 4330–4336.
26. Fokkema IF, Taschner PE, Schaafsma GC, et al. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 2011; 32: 557–563.
27. Bellissimo DB, Christopherson PA, Flood VH, et al. VWF mutations and new sequence variations identified in healthy controls are more frequent in the African-American population. Blood 2012; 119: 2135–2140.
28. Goodeve A, Eikenboom J, Castaman G, et al. Phenotype and genotype of a cohort of families historically diagnosed with type 1 von Willebrand disease in the European study, Molecular and Clinical Markers for the Diagnosis and Management of Type 1 von Willebrand Disease (MCMDM-1VWD). Blood 2007; 109: 112–121.
29. James PD, Notley C, Hegadorn C, et al. The mutational spectrum of type 1 von Willebrand disease: Results from a Canadian cohort study. Blood 2007; 109: 145–154.
30. Cumming A, Grundy P, Keeney S, et al. An investigation of the von Willebrand factor genotype in UK patients diagnosed to have type 1 von Willebrand disease. Thromb Haemost 2006; 96: 630–641.
31. Yadegari H, Driesen J, Pavlova A, et al. Mutation distribution in the von Willebrand factor gene related to the different von Willebrand disease (VWD) types in a cohort of VWD patients. Thromb Haemost 2012; 108: 662–671.
32. Federici AB, Bucciarelli P, Castaman G, et al. Management of inherited von Willebrand disease in Italy: results from the retrospective study on 1234 patients. Semin Thromb Haemost 2011; 37: 511–521.
33. Casari C, Lenting P, Wohner N, et al. Clearance of von Willebrand factor. J Thromb Haemost 2013; 11 (Suppl1): 202–11.
34. Smith NL, Chen MH, Dehghan A, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. Circulation 2010; 121: 1382–1392.
35. Federici AB. Acquired von Willebrand syndrome associated with hypothyroidism: a mild bleeding disorder to be further investigated. Semin Thromb Haemost 2011; 37: 35–40.
36. Federici AB, Budde U, Castaman G, et al. Current diagnostic and therapeutic approaches to patients with acquired von Willebrand syndrome: a 2013 update. Semin Thromb Haemost 2013; 39: 191–201.
37. Sanders YV, Giezenaar MA, Laros-van Gorkom BA, et al. von Willebrand disease and aging: an evolving phenotype. J Thromb Haemost 2014; 12: 1066–1075.
38. Montgomery RR. When it comes to von Willebrand disease, does 1 + 1 = 3? J Thromb Haemost 2006; 4: 2162–2163.
39. Casaña P, Martínez F, Haya S, et al. Q1311X: a novel nonsense mutation of putative ancient origin in the von Willebrand factor gene. Br J Haematol 2000; 111: 552–555.
40. Jorieux S, Gaucher C, Goudemand J, et al. A novel mutation in the D3 domain of von Willebrand factor markedly decreases its ability to bind factor VIII and affects its multimerisation. Blood 1998; 92: 4663–4670.
41. Johnsen JM, Auer PL, Morrison AC, et al. Common and rare von Willebrand factor (VWF) coding variants, VWF levels, and factor VIII levels in African Americans: the NHLBI Exome Sequencing Project. Blood 2013; 122: 590–597.
42. Haberichter SL, Balistreri M, Christopherson P, et al. Assay of the von Willebrand factor (VWF) propeptide to identify patients with type 1 von Willebrand disease with decreased VWF survival. Blood 2006; 108: 3344–3351.
43. ten Bosch JR, Grody WW. Keeping up with the next generation: massively parallel sequencing in clinical diagnostics. J Mol Diagn 2008; 10: 484–492.
44. Hampshire DJ, Goodeve AC. The international society on thrombosis and haematosis von Willebrand disease database: an update. Semin Thromb Haemost 2011; 37: 470–479.
45. Haberichter SL, Hille M, Morateck P, et al. Alanine Scanning Mutagenesis of the VWF-A1 Flanking Regions of VWF and the Impact on Ristocetin-induced VWF Interaction with Platelets. Blood 2013; 122: 1103.
46. Goodeve AC. VWF sequence variants: a data goldmine. Blood 2013; 122: 471–473.
47. Flood VH, Friedman KD, Gill JC, et al. No increase in bleeding identified in type 1 VWD subjects with D1472H sequence variation. Blood 2013; 121: 3742–3744.
48. Souto JC, Almasy L, Soria JM, et al. Genome-wide linkage analysis of von Willebrand factor plasma levels: results from the GAIT project. Thromb Haemost 2003; 89: 468–474.
49. Johnsen JM, Teschke M, Pavlidis P, et al. Selection on cis-regulatory variation at B4galnt2 and its influence on von Willebrand factor in house mice. Mol Biol Evol 2009; 26: 567–578.
50. Shavit JA, Manichaikul A, Lemmerhirt HL, et al. Modifiers of von Willebrand factor identified by natural variation in inbred strains of mice. Blood 2009; 114: 5368–5374.
51. Gill JC, Endres-Brooks J, Bauer PJ, et al. The effect of ABO blood group on the diagnosis of von Willebrand disease. Blood 1987; 69: 1691–1695.
52. Rydz N, Swystun LL, Notley C, et al. The C-type lectin receptor CLEC4M binds, internalizes, and clears von Willebrand factor and contributes to the variation in plasma von Willebrand factor levels. Blood 2013; 121: 5228–5237.

Material Suplementari Article 1

Supplementary Material to Batlle et al. "Molecular and clinical profile of von Willebrand disease in Spain (PCM–EVW–ES): Proposal for a new diagnostic paradigm" (Thromb Haemost 2015; 114.6)

Suppl. Material and Methods

Study design

A minor modification of the classification used was the inclusion of subtype 1H (historical) proposed by Montgomery et al[1] for cases in which an historical von Willebrand factor (VWF) level was confirmed but currently showing a very mild decrease or even a normal VWF plasma level. Despite not being recognized yet, this subclassification is appropriate as some patients with type 1 VWD show a progressive increase in VWF levels with age, but some still bleed despite having a normal VWF level, as shown in the Netherlands Win Study.[2] Type 1 diagnosis was restricted to patients with low levels of VWF at the recruitment moment of this study.

A minor additional modification is the inclusion of type 2A/2M which has been the subject of important controversy.[3]

Patients and controls

A 30 IU/dL VWF cut-off plasma level was used when a low VWF level was the only fulfilled criteria. There are several reasons for this: 1. The previous published view on type 1 VWD.[4] 2. The present work was not designed as an epidemiological study to establish either the prevalence of VWD or the distribution of its different types in Spain. Rather it was intended, as well as other objectives, to evaluate and validate the next generation sequencing (NGS) technology. 3. The rate of *VWF* mutations found in type 1 VWD decreases when VWF level >30 IU/dL.[5,6] 4. Due to financial and time constraints, to avoid submission (recruitment) of an overwhelming number of patients with borderline VWF levels.

Genetic analysis

DNA extraction. Genomic DNA was obtained from peripheral blood samples by using a QIAsymphony SP instrument and the QIAsymphony DNA Midi Kit (Qiagen). DNA concentration was adjusted to a range of 25 to 50 ng/µL.

Primer design. A total of 61 pairs of oligonucleotides have been designed (Accession NG_029217) capable of amplifying the relevant regions of the *VWF* gene (exons 1 to 52, intronic flanking regions and approximately 1300 bp of *VWF* promoter region). Primers used were designed taking into account the particular differences between the gene and pseudogene[7] to ensure highly specific amplification, as well as the absence of single nucleotide poylmorphisms (SNPs) in primer binding sequences that could result in preferential or single allele amplification, and lead to missing alleles (dbSNP [Build 138] and 1000 Genomes polymorphism database). Table S1 shows the primer sequences, locations in human genome sequence, and the polymerase chain reaction (PCR) product sizes.

PCR amplification. PCR sizes were adjusted to meet the required size criteria of the sequencing platform (between 376 and 450 bp bases/amplicon).

The Access Array™ System (Fluidigm) allows 2304 different PCR reactions by combining 48 samples with 48 pairs of primers (the 61 primer pairs were adjusted to 48 by creating 11 duplex and one triplex). The samples were loaded into the Access Array™ in combination with a distinctive short sequence identification label that acted as a barcode and which were incorporated into each set of amplicons from the same sample during the amplification.

Suppl. Results

Genetic results

Mutation versus variant

Indeed, genetic nomenclature has been and remains controversial in many ways and it is important to specify in each case the exact meaning of the terms used. We have attempted to use the term ‘mutation’ to indicate ‘a disease-causing change’. Conversely, we have used the word ‘variant’ for alterations with an unknown functional effect, or ‘variants of unknown significance’ (VUS). Similarly, the term ‘polymorphism’ was used both to indicate ‘a non disease-causing change’ and ‘a change found at a frequency of 1% or higher in any population included in the SNP databases’.

Types of mutations

Missense mutations. A total of 466 missense mutations (64% of all mutations) were detected (119 different) in 220 unrelated families. The top 10 mutations were found 194 times, and the five most frequent mutations (p.Arg1205His, p.Arg1374His, p.Pro2063Ser, p.Arg1374Cys, p.Arg854Gln), all of them described previously, represented 25% of missense mutations. The predicted impact of these mutations was

examined with two different *in silico* programs showing: 21 benign mutations, 20 possible damaging and 78 probably damaging for PolyPhen–2; 39 tolerated and 80 deleterious according to Sorting Intolerant From Tolerant (SIFT) analysis. Correlation in predictions between both algorithms was weak: only seven mutations (p.Pro2063Ser, p.Cys1899Arg, p.Cys1946Phe, p.Pro2646Thr, p.Cys2283Arg, p.Arg2118Trp, p.Cys2773Ser) were identified as non-deleterious (tolerated and benign) by both programs. None of these changes were included in the *VWF* mutation databases, except for the controversial p.Pro2063Ser, previously described both as a polymorphism and VWD causative mutation.

Null mutations. A total of 33 stop mutations (11 different) were found in 24 unrelated families, 20 patients were homozygous and 13 were heterozygous for stop mutations. The putative gene conversion mutation p.Gln1311Ter was found most frequently, in eight patients from six unrelated families.

Eighteen frameshift variants (10 different) were found in 15 unrelated families; one of these variants (c.376delT; p.Tyr126ThrfsTer49) was very prevalent in our population and was found in seven unrelated families. Remarkably, this mutation was always found with c.375G>C (synonymous change), c.4146G>T (synonymous change) and c.6197A>G (p.Asn2066Ser), indicating a founder effect. Furthermore, 12 in-frame mutations (five different, including three deletions and two insertions) were detected in six unrelated families.

Finally, among null mutations we detected two large deletions: one in the homozygous state comprising exons 16 to 43 detected by consecutive regions of zero coverage in NGS and then confirmed by multiplex ligation-dependent probe amplification (MLPA), and one heterozygous, including exons 11 to 32 and not detected by NGS, but which was then discovered by MLPA.

Variants in splice site regions. A total of 26 splice site variants (12 different) were found in 19 families (7%) as a putative causative mutation. Nine of these mutations comprised the consensus 5′ GT donor splice site or the 3′ AG acceptor splice site of several introns (5, 10, 13, 25, 31, 36, 37, 41 and 45) and would, therefore, have an undisputed effect on *VWF* mRNA processing and patient phenotype. The remaining three potential splice site mutations detected (c.7730-4C>G, c.8254-5T>G and c.8155+3G>C) encompassed neighboring consensus nucleotides. *In silico* analysis was able to predict the score variation of all the affected splice sites in introns 45, 50 and 51.[8]

Synonymous, intronic and upstream mutations. Besides mutations that have a very evident effect (stop, frameshift, acceptor or donor splice site, etc.) and missense mutations, whose effects are more variable, we have kept on the Registry all those

variants in introns, promoter and synonym changes whose frequency in the general population and in different populations studied in the project 1000 Genomes was below 1%. In fact for most of these variants the frequency was close to zero. Although it is hard to demonstrate, any nucleotide modification could be considered a potential candidate to alter the splicing, especially when an intronic or synonymous mutation appears next to an intron/exon junction.[9,10] We found a total of 166 (81 different) mutations in this group (for details see Figure S1). However, in only 20 patients from 13 families was one such mutation found alone as a putative exclusively responsible mutation. In such cases in particular it will be necessary to perform additional studies in order to demonstrate the underlying contribution of the mutation to clinical manifestations.

Suppl. Table 1. Primers used to amplify the *VWF*. Primers in black were designed by Fluidigm custom-designed assays and primers in green are those proposed in order to avoid the missed regions due to the proximity of the primers to the coding regions. SNPs positions have been taken into consideration when designing the primers and *in silico* PCR has been developed. The shadowed area contains specific primers designed to amplify the *VWF* in the pseudogene region (exons 23 to 34). The nucleotides outlined in red correspond to positions that are specific (different) for the gene in respect of the pseudogene. Forward primers are linked in 5′ to sequence ACACTGACGACATGGTTCTACA and reverse primers are linked in 5′ to sequence TACGGTAGCAGAGACTTGGTCT. Such universal tails are necessary to construct libraries with adapters specific for the MiSeq desktop sequencer and add 'the molecular barcodes' concomitantly with VWF amplification.

Name	Amplifiedregion	Forward primer (5′-3′)	Reverse primer (5′-3′)	PCR size (bp)	Genomeregion (hg19)
VWP1	Promoter	CCCCAAAACAAGTGGAGACAGG	GCCATCGGGATTTCCACTTCCA	434	chr12:6235188-6235621
VWP2		AAGAACATTTGTTGTATTTAGCACAG	GAGACAGTAGAAAAGGCCATTGG	446	chr12:6234855-6235300
VWP3		TGATCCCTTGACCTCGTGATCC	CTTCTTCCCAGCCTCCCATTCC	448	chr12:6234511-6234958
VWP4		TGCCCATTGCATCAGTTACTTATTTT	CCAGGGAAGTTGAGAAAAACACC	449	chr12:6234154-6234602
VWP5		CCAAATGGACTGTTTTTGTCAGGG	CAGGGGATTGGCCTCCTTTTAAT	446	chr12:6233844-6234289
VWF_E1	E1	AGGTCTCCCTCCAAACTCTACAA	GGATCAGTCAGTCCTGCATCTTC	424	chr12:6233525-6233948
VWF_E2	E2	GGGGTCACGAGGATCAATCTTTT	TCTGTGCCACCTTTATGCTTCTT	425	chr12:6232070-6232494
VWF_E3	E3	CACAGCCCAGTTTCTATCAGAGG	AAAGGCTGTTCCTCTCCTTGTTC	419	chr12:6230201-6230619
VWF_E4	E4	AAGACTTTTTGGGGCGTTTTCTG	TCTAACCCCAACCCCATGTTGTA	438	chr12:6219814-6220251
VWF_E5	E5	TTTGCACCATGTTCTGAACCCTA	CCAGGGAAGGCATGTTAGTGAAG	407	chr12:6219451-6219857
VWF_E6	E6	CAAAACCACCAGCAGACCTAGAA	ACTACACCAAGTCATTAGCCCAG	425	chr12:6204415-6204839
VWF_E7	E7	CAGGGCTAAGTCTCAGTGCC	CGTGACACAGCCCCGAAGCACC	383	chr12:6184410-6184792
VWF_E8	E8	GTAAGGGCCTCACAAGATGGAAG	TTCATCTCACTTCCCAAAGCCAA	406	chr12:6182713-6183118
VWF_E9	E9	GAACAAGTTCTTTGAGCTTCCTGG	CATTGCCTGCCACACCTGTCC	401	chr12:6181328-6181728
VWF_E10	E10	TTCGTTTGGGGACTGTGATAACT	GGCCTTACAGCTCATCAACCTTA	448	chr12:6180162-6180609
VWF_E11	E11	TTAGACTGGTTTGGGCAAAGGAC	CTGGGTTTCTGGATGAATGAGGA	434	chr12:6174141-6174574
VWF_E12	E12	TCTCTGATTAAGAGGGTCCTGGG	CCTTCTCTAGTCTTCCTCCTCCA	429	chr12:6173203-6173631
VWF_E13	E13	GCTCTTCAGTGCTACCATCCTTT	GAAGGGGCATCCAGAAAACAAAC	401	chr12:6171934-6172334
VWF_E14	E14	CTAAACAACTATGCCGCTGCTTT	CCCAGTTCAGATGATCAAGTGCT	424	chr12:6166849-6167272

Name	Amplifiedregion	Forward primer (5′-3′)	Reverse primer (5′-3′)	PCR size (bp)	Genomeregion (hg19)
VWF_E15	E15	GCACTGGGCTATTTCCAGGGG	CCTACGCCCTCTTTCCACAGG	416	chr12:6165915-6166330
VWF_E16	E16	CACAGCTACAAGGGGTGGCAAG	CATCCATGAAGTAAAGGACTTGGG	416	chr12:6161622-6162037
VWF_E17	E17	CGTTAGCAAGCTGTGCTTCAGG	AAACCCAGAAATGAAGGCGATCC	391	chr12:6155734-6156124
VWF_E18	E18	AGAAGAGCTGCCACCAATGTAAG	CACTCATCCCTGCCTACAAGAAA	447	chr12:6153395-6153841
VWF_E19	E19	CTTACCCGTAGGCTCAAGTCTCA	AGGATGGACACAGGTGATGAAAA	402	chr12:6145362-6145763
VWF_E20	E20	CTGTGTTCCTTCATTGCCTCCAT	CAGAGGGAAAATGGACAACCACT	427	chr12:6143653-6144079
VWF_E21	E21	CCAATCTTCTGGTCTGGTGAGAG	GCACTGTCGCTAAAATGGGATTG	424	chr12:6140388-6140811
VWF_E22	E22	CTGTTACCGTCTCTTGGTCATGG	AGGGAGCAGAAAACACTCCAAAG	446	chr12:6138446-6138891
VWF_E23	E23	TGTCCCGCTCCTCCTCCTCTT	TCCCACCCTGCAGCCACCTGG	431	chr12:6134984-6135414
VWF_E24	E24	ATGTCCCTCTGTCCCCATT**G**T**C**	GGTGCAGATCATGTCAAGACAC**G**	395	chr12:6134628-6135022
VWF_E25	E25	CTCCTTCTACCTG**GA**CCCCC**T**	GCCATCCAGTCCCTACTAA**CA**C**T**	409	chr12:6132665-6133073
VWF_E26	E26	CAACATAGCAAGACCCCATCT**G**	ACCTTTCCATCCATCCCTATCC	448	chr12:6131846-6132293
VWF_E27	E27	GAGGGATCAATCAAGGCACAAG**C**	CTTTTTACCCAAAACCTAGTCTCT**A**	410	chr12:6130980-6131389
VWF_E28a	E28	TCAGAAGTGTCCA**C**AGGTTCTT**C**	GCATACTTCACCTGGCTGGCAA**T**	433	chr12:6128535-6128967
VWF_E28b		**C**AGAAGTGGGTCCGCGTGGC**C**	CACAGAGGTAGCTAACGATCTC**G**	445	chr12:6128209-6128653
VWF_E28c		CCTCAAGCAGATCCGCCTCAT**C**	GCCTCGCTGAAGGGGTACTCC	402	chr12:6127920-6128321
VWF_E28d		GGC**G**TT**C**GTCC**T**GGAAGGAT**C**	CACCTGGATGTCTCCAGGCA**G**	379	chr12:6127709-6128087
VWF_E28e		CAGGGTGACCGGGAGCAGG**C**	CCAGGATTAGAACCCGAGTC**G**	376	chr12:6127429-6127804
VWF_E29	E29	TTGTAGGCCTGG**T**GGCCA**T**TG**T**	CTGCATCCAGCCTGTGGCACC	450	chr12:6125653-6126102
VWF_E30	E30	GGTGACGAAGAGGCTCTTTTTG**T**	CTTAAAAGCTGAATGATTCAGA**A**	413	chr12:6125479-6125891
VWF_E31	E31	CCACCGT**T**AAGACAGGGTGTC**G**	TTT**G**GTTTCCTAATCACATCGTG**G**	395	chr12:6125072-6125466
VWF_E32	E32	GAACATCTTCCTCATAGGGCTG**A**	CACCCTGGGGTCTCTTGAATA**C**	402	chr12:6122551-6122952
VWF_E33	E33	CACCTCAGCC**T**CATGTCCCTA**T**	ACTGACCCCTAGAATTGAACAC**A**	430	chr12:6121034-6121463
VWF_E34	E34	GCCTCCTTGCTGTGTAGGCC**T**	AGGGTACTTCTGGGCTGGTGG	398	chr12:6120679-6121076
VWF_E35	E35	TTTGGATAACGTCGCATCCATCC	GTGGTACTCCTCTCCTCCACTTA	401	chr12:6105030-6105430
VWF_E36	E36	AGTGTGATGAATGTGCAGGAACT	GCCATTCTTGATCCTCACCAGAA	427	chr12:6103443-6103869
VWF_E37	E37	CCAGGGCCACTCAGTTTATCTTT	GGCTGAGCAAGCCCAGTTAG	447	chr12:6102969-6103415
VWF_E38	E38	GTTGAATCAGCTGTGCCCATTTC	CCCTATCCCTAATGATCCCCGTA	406	chr12:6100851-6101256

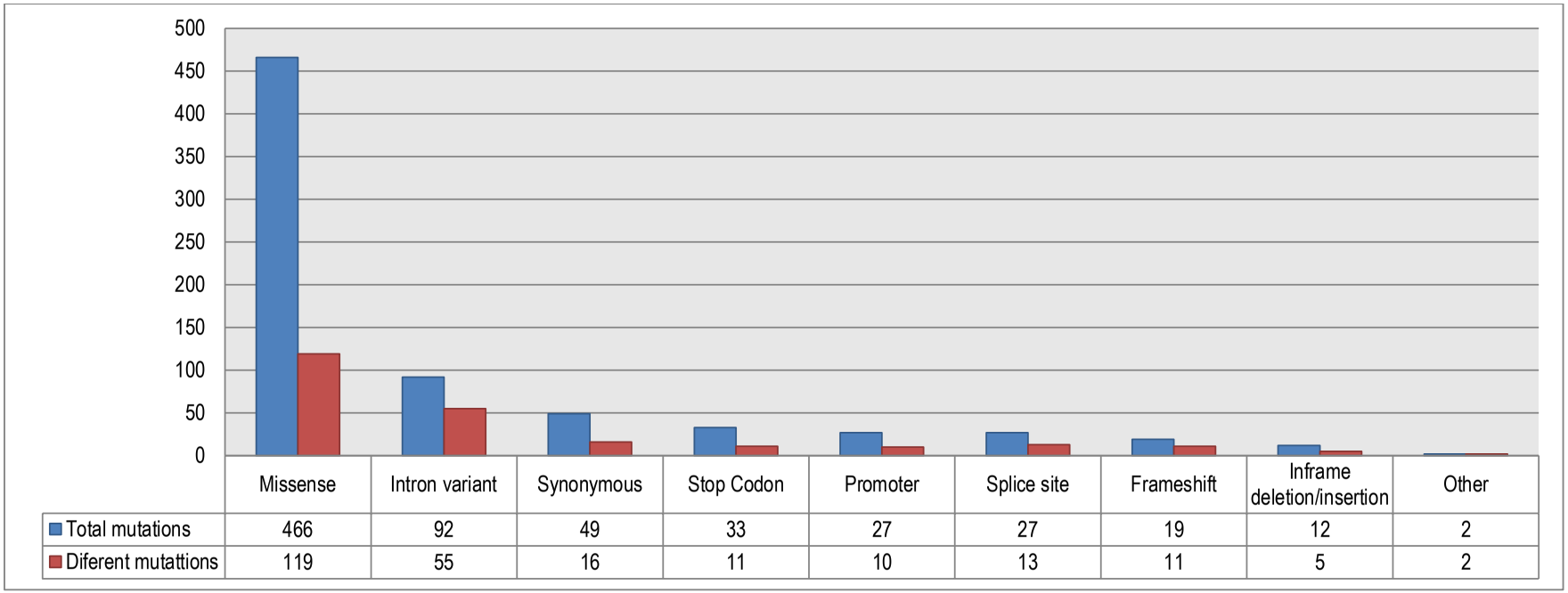
Name	Amplifiedregion	Forward primer (5′-3′)	Reverse primer (5′-3′)	PCR size (bp)	Genomeregion (hg19)
VWF_E39	E39	GGCTTCGTACCTAGAATGTCCTG	ACTCTTAAGTCTGACCGTTGCTG	404	chr12:6094481-6094884
VWF_E40	E40	GAGGGTAAATGTGAGGGTGTCAA	TTGGAGTATCCAGTCGGTCCTTA	405	chr12:6094070-6094474
VWF_E41	E41	GTAGCATCCCACTCACAGGTAAT	CAGCTGCTTCTTCATAAACCCAA	439	chr12:6092058-6092496
VWF_E42	E42	GGCACCCTATAGCATAGCTGAAT	ATGCTGGGTAGGATGCAAATCTT	447	chr12:6090788-6091234
VWF_E43	E43	CCCTGAGAGACTGGTCATAGCTT	CCTTCCTAAGATGCCCTCCTCTA	442	chr12:6085205-6085646
VWF_E44	E44	TGCAGCTGATGTAAGACTTCGTT	AAAGGGCAGGGAATGAGATGAAA	409	chr12:6080672-6081080
VWF_E45	E45	CCAGGCTTCACGTCTAGAAACC	GGTCCTATCCATTTCCCTAAGTTG	405	chr12:6078262-6078666
VWF_E46	E46	ATCCTCCCATTCATCCCATCTCT	GCTTTACAATGACTTGCCTGCTC	415	chr12:6077205-6077619
VWF_E47	E47	GATGGCTCATGGGAGCTATGG	TATTTCTGAGGCATGCAGTTTGG	408	chr12:6076471-6076878
VWF_E48	E48	GGCTGCTACAACTTGAGTGAGAT	CTGAACCAAACTTAGCCCCTCTT	417	chr12:6062564-6062980
VWF_E49	E49	GAGTCTACATAGCAGCCCTGTTC	GAGCTGGGATGCACACTATTCAG	427	chr12:6061477-6061903
VWF_E50	E50	TTACATCTGGGCACTTAAGCACA	GGAAAGAGATTGCTCCATGGGAT	414	chr12:6060757-6061170
VWF_E51	E51	GGCTGGCTTTATTTGGTTACTGTG	GGATTTGCTGAGTCTTTCCTCCA	440	chr12:6058733-6059172
VWF_E52	E52	AGATCAGACCTGCCTTGCTTG	AAGAGCTCAGCCTTTATTGTGGG	433	chr12:6058054-6058486

Suppl. Table 2. *F8* mutations validated by *Sanger* in patients of the Registry with a final diagnosis of hemophilia A and carriers.

Patient	Genotype	Consequence	Exon	Sift	PolyPhen	HGVSc	HGVSp	Described
*1 (hemophilic)	-	-	-	-	-	No mut	-	-
2 (hemophilic)	hem	missense_variant	15/26	deleterious(0.02)	probably_damaging(1)	c.5305G>A	p.Gly1769Arg	YES
3 (hemophilic)	hem	missense_variant	14/26	deleterious(0)	probably_damaging(0.997)	c.5123G>A	p.Arg1708His	YES
4 (hemophilic)	hem	missense_variant	9/26	deleterious(0.01)	probably_damaging(0.952)	c.1339T>C	p.Phe447Leu	NO
5 (hemophilic)	hem	missense_variant	9/26	deleterious(0.01)	probably_damaging(0.952)	c.1339T>C	p.Phe447Leu	NO
6 (hemophilic)	hem	missense_variant	9/26	deleterious(0.01)	probably_damaging(0.952)	c.1339T>C	p.Phe447Leu	NO
7 (hemophilic)	hem	missense_variant	18/26	deleterious(0.04)	probably_damaging(0.999)	c.5954G>A	p.Arg1985Gln	YES
8 (hemophilic)	hem	missense_variant	26/26	deleterious(0)	probably_damaging(0.999)	c.6968G>A	p.Arg2323His	YES
	hem	missense_variant	14/26	tolerated(0.19)	benign(0)	c.4264T>C	p.Tyr1422His	NO
9 (carrier)	het	missense_variant	15/26	deleterious(0.02)	probably_damaging(1)	c.5305G>A	p.Gly1769Arg	YES
10 (carrier)	het	missense_variant	14/26	deleterious(0)	probably_damaging(0.997)	c.5123G>A	p.Arg1708His	YES
11 (carrier)	het	synonymous_variant	2/26			c.222G>A	c.222G>A(p.=)	NO
12 (carrier)	het	missense_variant	9/26	deleterious(0.01)	probably_damaging(0.952)	c.1339T>C	p.Phe447Leu	NO
13 (carrier)	het	missense_variant	18/26	deleterious(0.04)	probably_damaging(0.999)	c.5954G>A	p.Arg1985Gln	YES

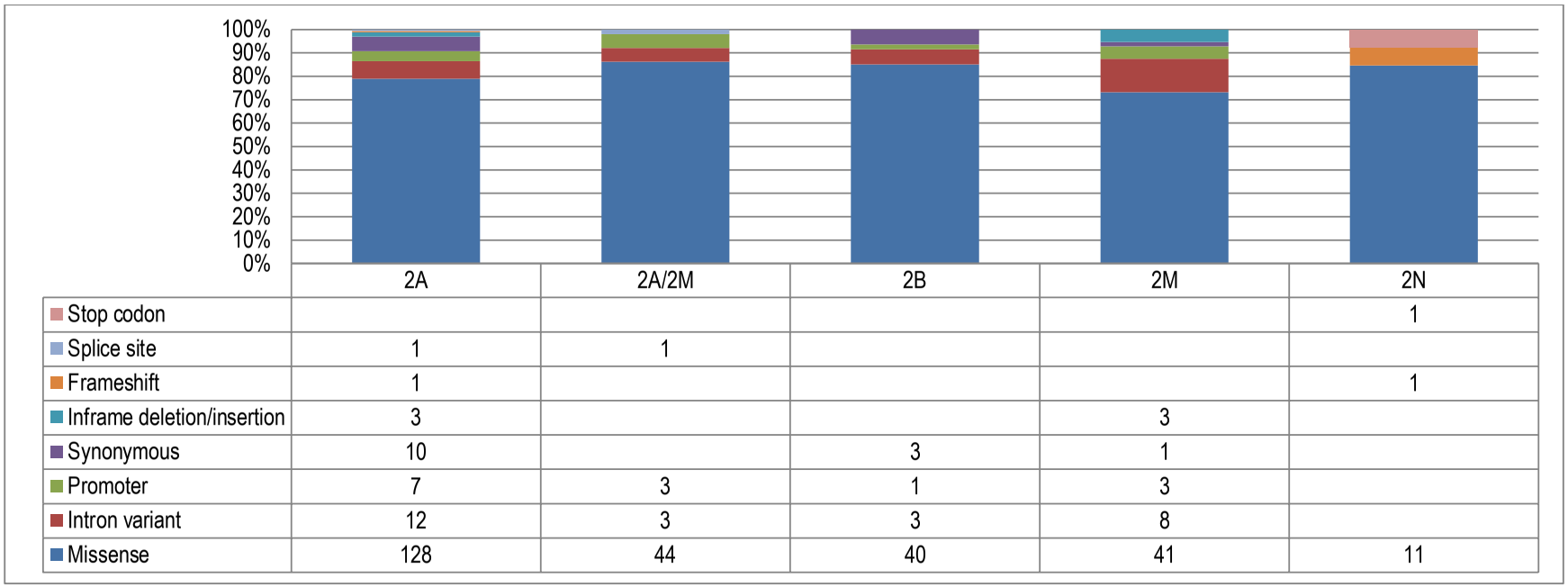
** Possible error in sampling. Pending of the study in a new sample.*

Suppl. Figure 1: Mutations (total and different) categorized by consequence found in patients included in the PCM–EVW–ES.



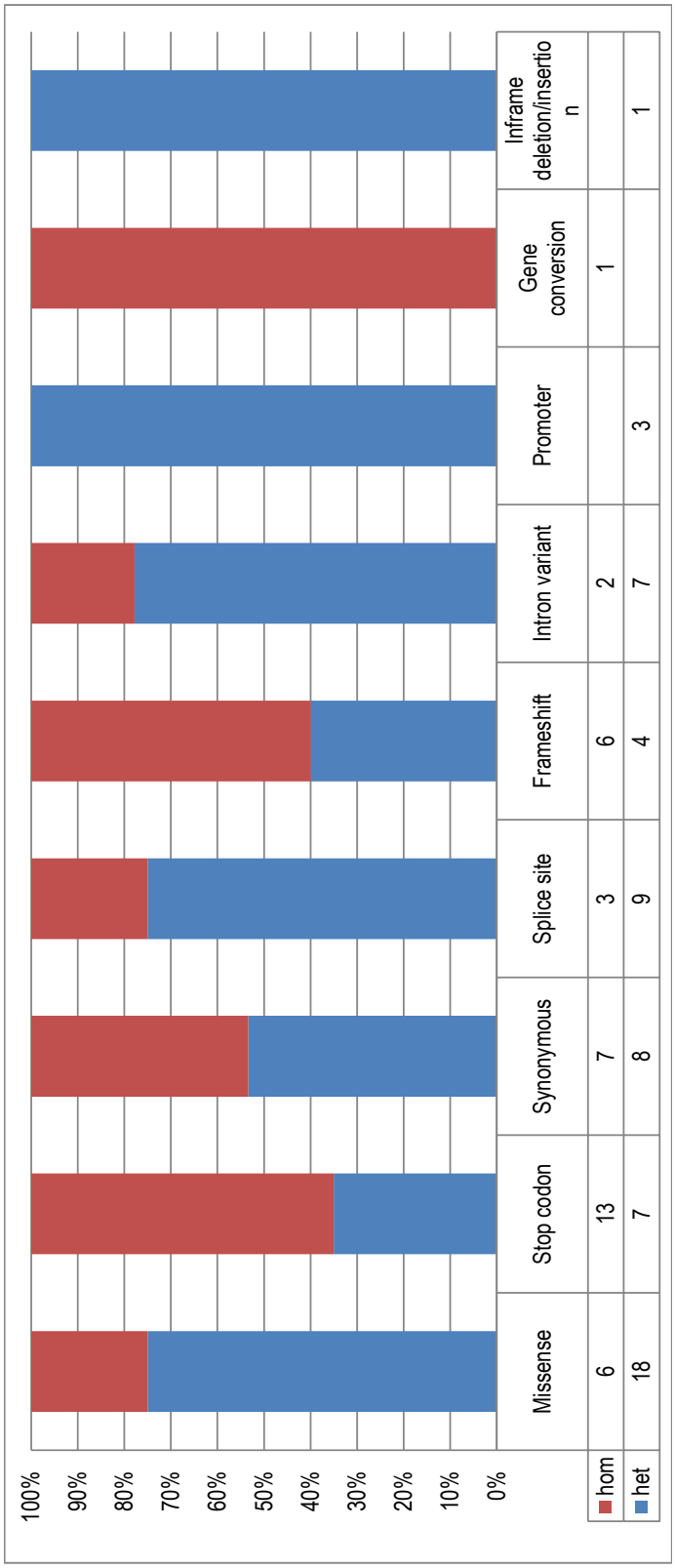
	Missense	Intron variant	Synonymous	Stop Codon	Promoter	Splice site	Frameshift	Inframe deletion/insertion	Other
Total mutations	466	92	49	33	27	27	19	12	2
Diferent mutattions	119	55	16	11	10	13	11	5	2

Suppl. Figure 2. Distribution of changes by type in qualitative defects.



	2A	2A/2M	2B	2M	2N
Stop codon					1
Splice site	1	1			
Frameshift	1				1
Inframe deletion/insertion	3			3	
Synonymous	10		3	1	
Promoter	7	3	1	3	
Intron variant	12	3	3	8	
Missense	128	44	40	41	11

Suppl. Figure 3. Distribution of changes by consequence and genotype state in type 3 VWD patients. Hom, homozygous; het, heterozygous.



Appendix: study group members

The members of the PCM-EVW-ES Study Group are:

Javier Batlle, A Coruña, Spain; Almudena Pérez-Rodríguez, A Coruña, Spain; Irene Corrales, Barcelona, Spain; María Fernanda López-Fernández, A Coruña, Spain; Ángela Rodríguez-Trillo, A Coruña, Spain; Esther Lourés, A Coruña, Spain; Ana Rosa Cid, Valencia, Spain; Santiago Bonanad, Valencia, Spain; Noelia Cabrera, Valencia, Spain; Andrés Moret, Valencia, Spain; Rafael Parra, Barcelona, Spain; María Eva Mingot-Castellano, Málaga, Spain; Ignacia Balda, Las Palmas de Gran Canaria, Spain; Carmen Altisent, Barcelona, Spain; Rocío Pérez-Montes, Santander, Spain; Rosa María Fisac, Segovia, Spain; Gemma Iruín, Vizcaya, Spain; Blanca Pinedo, Guadalajara, Spain; Sonia Herrero, Guadalajara, Spain; Inmaculada Soto, Asturias, Spain; Ángeles Rodríguez, Asturias, Spain; Nuria Fernández Mosteirín, Zaragoza, Spain; Beatriz de Rueda, Zaragoza, Spain; Víctor Jiménez-Yuste, Madrid, Spain; Mª Teresa Álvarez, Madrid, Spain; Nieves Alonso, Badajoz, Spain; Dolores Vilariño, Santiago de Compostela, Spain; Olga Arija, Lugo, Spain; Rosa Campos, Cádiz, Spain; María José Paloma, Navarra, Spain; Nuria Bermejo, Cáceres, Spain; Teresa Toll, Barcelona, Spain; José Mateo, Barcelona, Spain; Karmele Arribalzaga, Madrid, Spain; Pascual Marco, Alicante, Spain; Ángeles Palomo, Málaga, Spain; María Ángeles Martín, Madrid, Spain; Lizheidy Sarmiento Serrato, Madrid, Spain; María Paz Maluenda, Madrid, Spain; Belén Iñigo, Madrid, Spain; María del Mar Nieto, Jaén, Spain; Rosa Vidal, Madrid, Spain; María Paz Martínez, Ávila, Spain; Reyes Aguinaco, Tarragona, Spain; Jesús María César(†), Madrid, Spain; María Ferreiro, Pontevedra, Spain; Javier García-Frade, Valladolid, Spain; Ana María Rodríguez-Huerta, Madrid, Spain; Jorge Cuesta, Toledo, Spain; Ramón Rodríguez-González, Madrid, Spain; Faustino García-Candel, Murcia, Spain; Rosa Cornudella, Zaragoza, Spain; Carlos Aguilar, Soria, Spain; Nina Borràs, Barcelona, Spain; and Francisco Vidal, Barcelona, Spain.

(†) Jesús M. César died on August 25th, 2014

Suppl. References

1. Montgomery RR, Christopherson P, Bellissimo DB, et al. The complete type I VWD cohort of the Zimmerman Program for the Molecular and Clinical Biology of VWD-Phenotypic assignment, mutation frequency, and bleeding assessment. *Blood*. 2013;122(21):332-332.
2. Sanders YV, Giezenaar MA, Laros-van Gorkom BA, et al. von Willebrand disease and aging: an evolving phenotype. *J Thromb Haemost*. 2014;12(7):1066-1075.
3. Batlle J, Pérez-Rodríguez A, Franqueira MD, López-Fernández MF. Type 2M von Willebrand disease: a variant of type 2A? *J Thromb Haemost*. 2008;6(2):388-390.
4. Sadler JE. Von Willebrand disease type 1: a diagnosis in search of a disease. *Blood*. 2003;101(6):2089-2093.
5. Tosetto A, Rodeghiero F, Castaman G, et al. A quantitative analysis of bleeding symptoms in type 1 von Willebrand disease: results from a multicenter European study (MCMDM-1 VWD). *J Thromb Haemost*. 2006;4(4):766-773.
6. Goodeve A, Eikenboom J, Castaman G, et al. Phenotype and genotype of a cohort of families historically diagnosed with type 1 von Willebrand disease in the European study, Molecular and Clinical Markers for the Diagnosis and Management of Type 1 von Willebrand Disease (MCMDM-1VWD). *Blood*. 2007;109(1):112-121.
7. Corrales I, Ramírez L, Altisent C, Parra R, Vidal F. Rapid molecular diagnosis of von Willebrand disease by direct sequencing. Detection of 12 novel putative mutations in VWF gene. *Thromb Haemost*. 2009;101(3):570-576.
8. Corrales I, Ramírez L, Altisent C, Parra R, Vidal F. The study of the effect of splicing mutations in von Willebrand factor using RNA isolated from patients' platelets and leukocytes. *J Thromb Haemost*. 2011;9(4):679-688.
9. Zatkova A, Messiaen L, Vandenbroucke I, et al. Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. *Hum Mutat*. 2004;24(6):491-501.
10. Dehainault C, Michaux D, Pagès-Berhouet S, et al. A deep intronic mutation in the RB1 gene leads to intronic sequence exonisation. *Eur J Hum Genet*. 2007;15(4):473-477.

ARTICLE 2

Títol: "Perfil molecular i clínic de la malaltia de von Willebrand a Espanya (PCM-EVW-ES): extens estudi genètic de 480 pacients mitjançant seqüenciació massiva"

Resum

El diagnòstic molecular de pacients amb la malaltia de von Willebrand (VWD) està pendent en la majoria de les poblacions a causa de la complexitat i l'alt cost de les anàlisis moleculars convencionals. La necessitat de la caracterització molecular i clínica de la VWD a Espanya va impulsar la creació d'un projecte multicèntric (PCM-EVW-ES) que va donar com a resultat l'estudi prospectiu de la cohort més gran de pacients amb tots els tipus de VWD. L'anàlisi molecular de les regions rellevants del *VWF*, incloent les regions intròniques i el promotor, es va assolir en els 556 individus reclutats gràcies al desenvolupament d'un protocol simple, innovador i de baix cost, basat en la tecnologia de microfluids i la seqüenciació massiva. Es van identificar un total de 704 variants (237 diferents) al llarg de *VWF*, de les quals 155 no s'havien registrat prèviament a la base de dades internacional de mutacions. El potencial efecte patogènic d'aquestes variants es va avaluar mitjançant programes d'anàlisi *in silico*. A més, es van analitzar 4 repeticions en tàndem per avaluar l'origen ancestral de les mutacions recurrents. El resultat de l'anàlisi genètica va permetre la reclassificació de 110 pacients, la identificació de 37 portadors asimptomàtics (importants per a l'assessorament genètic) i la reincorporació al registre de 43 pacients exclosos anteriorment pels resultats de fenotip. En total, 480 pacients van ser diagnosticats definitivament. Es van identificar mutacions candidates en tots els pacients, excepte en 13 amb VWD de tipus 1, donant lloc a una alta correlació genotip-fenotip. Les nostres dades reforcen la importància i la utilitat de la genètica en el diagnòstic de la VWD. La implementació progressiva de l'estudi molecular com a prova de primera línia pel diagnòstic de rutina d'aquesta patologia donarà lloc a una atenció cada vegada més personalitzada i eficaç per a aquest pacients.

Referència

Borràs N, Batlle J, Pérez-Rodríguez A, López-Fernández MF, Rodríguez-Trillo Á, Lourés E, Cid AR, Bonanad S, Cabrera N, Moret A, Parra R, Mingot-Castellano ME, Balda I, Altisent C, Pérez-Montes R, Fisac RM, Iruín G, Herrero S, Soto I, de Rueda B, Jiménez-Yuste V, Alonso N, Vilariño D, Arija O, Campos R, Paloma MJ, Bermejo N, Berrueco R, Mateo J, Arribalzaga K, Marco P, Palomo Á, Sarmiento L, Iñigo B, Nieto MDM, Vidal R, Martínez MP, Aguinaco R, César JM, Ferreiro M, García-Frade J, Rodríguez-Huerta AM, Cuesta J, Rodríguez-González R, García-Candel F, Cornudella R, Aguilar C, Vidal F, Corrales I.

Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): comprehensive genetic analysis by next-generation sequencing of 480 patients.

Haematologica. 2017 Dec; 102(12):2005-2014. doi:10.3324/haematol.2017.168765. PMID: 28971901.

Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): comprehensive genetic analysis by next-generation sequencing of 480 patients

Nina Borràs,[1,2] Javier Batlle,[3] Almudena Pérez-Rodríguez,[3] María Fernanda López-Fernández,[3] Ángela Rodríguez-Trillo,[3] Esther Lourés,[3] Ana Rosa Cid,[4] Santiago Bonanad,[4] Noelia Cabrera,[4] Andrés Moret,[4] Rafael Parra,[1,5] María Eva Mingot-Castellano,[6] Ignacia Balda,[7] Carme Altisent,[5] Rocío Pérez-Montes,[8] Rosa María Fisac,[9] Gemma Iruín,[10] Sonia Herrero,[11] Inmaculada Soto,[12] Beatriz de Rueda,[13] Víctor Jiménez-Yuste,[14] Nieves Alonso,[15] Dolores Vilariño,[16] Olga Arija,[17] Rosa Campos,[18] María José Paloma,[19] Nuria Bermejo,[20] Rubén Berrueco,[21] José Mateo,[22] Karmele Arribalzaga,[23] Pascual Marco,[24] Ángeles Palomo,[6] Lizheidy Sarmiento,[25] Belén Iñigo,[26] María del Mar Nieto,[27] Rosa Vidal,[28] María Paz Martínez,[29] Reyes Aguinaco,[30] Jesús María César,[31] María Ferreiro,[32] Javier García-Frade,[33] Ana María Rodríguez-Huerta,[34] Jorge Cuesta,[35] Ramón Rodríguez-González,[36] Faustino García-Candel,[37] Rosa Cornudella,[38] Carlos Aguilar,[39] Francisco Vidal*[1,2,40] and Irene Corrales*[1,2]

[1]Banc de Sang i Teixits, Barcelona; [2]Vall d'Hebron Research Institute, Universitat Autònoma de Barcelona (VHIR-UAB); [3]Complexo Hospitalario Universitario A Coruña, INIBIC; [4]Hospital Universitario y Politécnico La Fe, Valencia; [5]Hospital Universitari Vall d'Hebron, Barcelona; [6]Hospital Regional Universitario de Málaga; [7]Hospital Universitario Dr. Negrín, Las Palmas de Gran Canaria; [8]Hospital Universitario Marqués de Valdecilla, Santander; [9]Salud Castilla y León, Segovia; [10]Hospital Universitario Cruces, Barakaldo; [11]Hospital Universitario de Guadalajara; [12]Hospital Universitario Central de Asturias, Oviedo; [13]Hospital Universitario Miguel Servet, Zaragoza; [14]Hospital Universitario La Paz, Madrid; [15]Hospital Infanta Cristina, Badajoz; [16]Complejo Hospitalario Universitario Santiago de Compostela; [17]Hospital Universitario Lucus Augusti, Lugo; [18]Hospital Jerez de la Frontera, Cádiz; [19]Hospital Virgen del Camino, Pamplona; [20]Hospital San Pedro de Alcántara, Cáceres; [21]Hospital Sant Joan de Deu, Barcelona; [22]Hospital de la Santa Creu i Sant Pau, Barcelona; [23]Hospital Universitario Fundación de Alcorcón, Madrid; [24]Hospital General de Alicante; [25]Hospital Universitario 12 de Octubre, Madrid; [26]Hospital Clínico San Carlos, Madrid; [27]Complejo Hospitalario de Jaén; [28]Fundación Jiménez Díaz, Madrid; [29]Hospital Nuestra Señora de Sonsoles, Ávila; [30]Hospital Universitario de Tarragona Joan XXIII; [31]Hospital Ramón y Cajal, Madrid; [32]Hospital Montecelo, Pontevedra; [33]Hospital del Río Hortega, Valladolid; [34]Hospital Gregorio Marañón, Madrid; [35]Hospital Virgen de la Salud, Toledo; [36]Hospital Severo Ochoa, Madrid; [37]Hospital Clínico Universitario Virgen de la Arrixaca, Murcia; [38]Hospital Clínico Universitario Lozano Blesa, Zaragoza; [39]Hospital Santa Bárbara, Soria and [40]CIBER de Enfermedades Cardiovasculares, Spain

*IC and FV contributed equally to this work

Haematologica 2017
Volume 102(12):2005-2014

Correspondence:
icorrales@bst.cat or fvidal@bst.cat

Received: March 16, 2017.
Accepted: September 20, 2017.
Pre-published: September 29, 2017.

doi:10.3324/haematol.2017.168765

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/102/12/2005

ABSTRACT

Molecular diagnosis of patients with von Willebrand disease is pending in most populations due to the complexity and high cost of conventional molecular analyses. The need for molecular and clinical characterization of von Willebrand disease in Spain prompted the creation of a multicenter project (PCM-EVW-ES) that resulted in the largest prospective cohort study of patients with all types of von Willebrand disease. Molecular analysis of relevant regions of the *VWF*, including intronic and promoter regions, was achieved in the 556 individuals recruited *via* the development of a simple, innovative, relatively low-cost protocol based on microfluidic technology and next-generation sequencing. A total of 704 variants (237 different) were identified along *VWF*, 155 of which had not been previously recorded in the international mutation database. The potential pathogenic effect of these variants was assessed by *in silico* analysis. Furthermore, four short tandem repeats were analyzed in order to evaluate the ancestral origin of recurrent mutations. The outcome of genetic analysis allowed for the reclassification of 110 patients, identification of 37 asymptomatic carriers (important for genetic counseling) and re-inclusion of 43 patients previously excluded by phenotyping results. In total, 480 patients were

definitively diagnosed. Candidate mutations were identified in all patients except 13 type 1 von Willebrand disease, yielding a high genotype-phenotype correlation. Our data reinforce the capital importance and usefulness of genetics in von Willebrand disease diagnostics. The progressive implementation of molecular study as the first-line test for routine diagnosis of this condition will lead to increasingly more personalized and effective care for this patient population. *clinicaltrials.gov Identifier: 02869074.*

Introduction

Von Willebrand disease (VWD) is the most common bleeding disorder, with a reported incidence of 0.01% to 1%.[1,2] The condition is caused by abnormalities related to the von Willebrand factor protein (VWF), which has an important role in primary hemostasis.[3,4] A diagnostic algorithm based on the patient's clinical history and the results of laboratory testing to assess VWF levels and functionality has enabled the classification of VWD into quantitative (type 1 and type 3) and qualitative (type 2) abnormalities.[1] The first-line tests used to classify VWD include the combination of VWF: antigen (Ag), VWF: ristocetin cofactor (RCo), factor VIII:C (FVIII:C), and the VWF:RCo/VWF:Ag ratio, whereas second-line tests include ristocetin-induced platelet aggregation (RIPA), VWF:FVIIIB, VWF: collagen binding (CB), VWF propeptide (pp), and multimer analysis, which are especially useful to differentiate between the type 2 VWD subtypes (2A, 2B, 2M, and 2N). It is essential to correctly identify these subtypes because their treatment differs. Nonetheless, this can be a challenging task as clinical and laboratory phenotypes are very heterogeneous,[5] and some tests, such as RIPA and multimeric analysis, are not available in all laboratories.[6] Furthermore, there is often a lack of clinical information, and the available laboratory tests are known to have intrinsic limitations.[1,7,8] In addition, sequential application of these analyses makes phenotypic VWD diagnosis increasingly more costly.

The VWF glycoprotein is encoded by *VWF* gene, a large (178 kb comprising 52 exons) and very polymorphic gene (>100 coding single-nucleotide polymorphisms (cSNPs)) with a highly homologous partial pseudogene (*VWFP*; exons 23-34),[9] characteristics that make specific amplification and sequencing of *VWF* difficult. Thus, molecular analysis of *VWF* is a challenge. Although genetic analysis is considered a valuable tool to support the diagnosis of VWD,[10,11] classic Sanger sequencing of the entire *VWF* is too costly for general use in all patients. Hence, this method has been applied, with slight variations, only to specific exons of *VWF*, depending on the VWD subtype. Exon 28 is sequenced for 2A, 2B, and 2M patients, and in type 2A, if a mutation is not found in exon 28, gene study is extended to exons 11-17 and 50-52. In 2N cases, exons 18 to 20 are sequenced, corresponding to the FVIII-binging site. Finally, analysis of the complete coding sequence is required in type 1 and type 3 patients, as potential mutations may be spread all along *VWF*.[12] Nonetheless, the advent of next-generation sequencing (NGS) technology is making molecular diagnosis of VWD by complete *VWF* sequencing faster and progressively less costly,[11] in particular when the new "desktop instruments" and simplified, optimized technologies for library preparation are used.[13]

Based on the results of a previous Spanish survey elucidating the difficulties of diagnosing VWD,[7] a multicenter, prospective project (PCM-EVW-ES, Molecular and Clinical Profile of von Willebrand Disease in Spain) was designed to centrally characterize a large multicenter cohort of VWD patients, with inclusion of NGS molecular analysis. As described in our previous report on this project,[14] this new technology enabled us to undertake a nationwide molecular epidemiologic study that supported a new scenario for VWD diagnostics, with genetic analysis being an indispensable first-line tool.

Herein, we provide an in-depth description of the technical aspects of NGS-based molecular characterization of *VWF* in the 556 individuals studied, complete details of the identified variants, and a description of the results of a genotypic-phenotypic correlation study in patients diagnosed with all types of VWD.

Methods

Patients

Molecular studies were performed on samples from 556 individuals with locally diagnosed VWD recruited in 38 Spanish hospitals participating in the PCM-EVW-ES project.[7,14] The inclusion criteria were one or more of the following: 1) VWF:Ag, VWF:RCo and/or VWF:CB ≤30% on at least two occasions regardless of the blood group, 2) detection of multimeric abnormalities, 3) evidence of decreased VWF:FVIIIB level in isolated FVIII deficiency, 4) presence of *VWF* candidate mutations, and 5) RIPA at a low ristocetin concentration.[14] Bleeding score (BS) and central laboratory phenotypic assessment are available in *Online Supplementary Methods*. The study was performed according to the Declaration of Helsinki, was approved by the local Research Ethics Committee, and all participants provided written informed consent.

VWF Access Array amplification and sequencing

We designed 61 pairs of oligonucleotides (Accession NG_029217) capable of amplifying exons 1 to 52, intronic flanking regions, and approximately 1300 bp of the promoter region that were described in our previous report.[14] The 48.48 Access Array integrated fluidic circuit (IFC; Fluidigm, San Francisco, CA, USA) is a nanofluidic chip that allows 2304 different polymerase chain reactions (PCRs) in a final volume of 30 nL by combining 48 samples with 48 primer pairs. As 61 primer pairs are needed for total *VWF* amplification, 12 multiplex PCRs were created, which allowed 2928 reactions per chip. The Access Array was processed following the manufacturers' recommendations (*Online Supplementary Methods* and *Online Supplementary Table S1*). The outcome was a pool of all *VWF* amplicons from the same sample, plus a distinctive short sequence identification label that acted as a barcode, incorporated into each set of PCRs during the amplification step. The final pools of up to 192 samples (four Access Array fusion libraries) were then combined and sequenced in a MiSeq platform (Illumina, San Diego, CA, USA; *Online Supplementary Figure S1*).

Data analysis and identification of genetic variants

Barcoded sequences were demultiplexed and analyzed individually. The NGS pipeline output, paired sequence files (FASTQ format), was used as input for analysis with the CLC Genomic Workbench software 8.0.2 (Qiagen, Hilden, Germany). After vari-

ant calling, the resulting files (VCF) were used as input for VariantStudio 2.2.1 (Illumina). The result of this step is the identification of potential pathogenic variants and filtering of the polymorphisms described to date in the SNP (dbSNP) and 1000 Genomes databases (Figure 1). The specific analytical parameters used are described in *Online Supplementary Methods*. Selected mutations detected were validated/confirmed by Sanger sequencing.[10] The nomenclature and criteria used to establish variant pathogenicity is also provided in *Online Supplementary Methods*.

In silico analysis

In silico prediction to evaluate the functional effects of putative pathogenic variants was performed using the Alamut Visual v.2.6.1 software (Interactive Biosoftware, Rouen, France). Missense prediction tools included PolyPhen-2, the Sorting Intolerant From Tolerant (SIFT), Mutation Taster, Mutation Assessor and Provean, whereas predictive tools for synonymous and splice site candidate mutations included GeneSplicer, Splicing Sequences Finder (SSF), Human Splicing Finder (HSF), Neural Network (NN) Splice and Maximum Entropy Modeling (MaxEnt). Upstream *VWF* variants were localized, visualized by the Integrative Genomics Viewer software, and compared in parallel with datasets from the Encyclopedia of DNA Elements (ENCODE) project to identify potential regulatory regions.[15]

Multiplex ligation-dependent probe amplification (MLPA)

VWF deletions/duplications were detected by MLPA using the SALSA MLPA P011 and P012 VWF kits (version B2; MRC–Holland, Amsterdam, The Netherlands), as previously described.[14,16] This method was applied to 5 patients whose genotype did not correlate with the phenotype

Microsatellite analysis

A multiplex fluorescent PCR described by our group,[17] comprising 3 *VWF* intragenic tetranucleotide short tandem repeats (STR-1, STR-2, and STR-3) and 1 dinucleotide repeat in the promoter region (WPA) was applied to genomic DNA samples.

Genotype-phenotype correlation

The correlation between genotype and phenotype was assessed by experts from the central laboratories of the PCM-EVW-ES who contrasted the results of the phenotypic test panel and the genetic analysis on the basis of the effect and localization of mutations and previous descriptions in the literature and/or databases.

Results

Nanofluidic *VWF* amplification and NGS output

The selected *VWF* regions were amplified with the Fluidigm Access Array. As each library is distinctively labeled, up to 384 samples can be pooled and sequenced in a single run. Fifteen Access Array and 5 MiSeq 500-cycle runs were needed to process the 556 samples (Figure 1, *Online Supplementary Results*). The output quality parameters obtained for each run and the resulting mean values are shown in *Online Supplementary Table S2*. In total, 25620 bp of *VWF* were sequenced in all samples, with homogeneous coverage for nearly all 61 amplicons (*Online Supplementary Figure S2*). Regions with low or no coverage were completed by Sanger sequencing.[10]

Genotype-phenotype correlation and classification

Of the 556 individuals recruited in the PCM-EVW-ES database, 442 had confirmed VWD based on central phenotype characterization. Mutation analysis by NGS was performed in all 556 individuals, and MLPA was additionally used in five (03012, 05011, 10011, 23010, 25005). Two large deletions were identified: c.1157_5620del in patient 03012 (type 2A) due to homologous recombination between short sequences of introns 10 and 32, and c.(1945+1_1946-1)_(7437+1_7438-1)del in the type 3 patient, 25005 (exact breakpoints pending characterization). Following mutation analysis, 105 were reclassified to a different subtype, five were excluded and 43 patients were reincluded (Figure 2) due to the presence of a candidate mutation in *VWF*: 19 type 3 VWD carriers, one type 2N VWD carrier, four type 1 VWD with borderline levels, nine type 1H VWD, three type 2M VWD with collagen binding mutations, and seven patients with uncertain classification. All these data, based on phenotyping and genotyping results, leads to a final diagnosis of VWD in 480 patients from 280 families (mean 1.7 members per family; range 1-21) who met the criteria for inclusion in the VWF registry (Figure 3). A total of 704 variants were identified, 237 were different, and 155 had not been described in the VWF European Association for Haemophilia and Allied Disorders (EAHAD) Coagulation Factor Variant Database compiled in the Leiden Open Variation Database (LOVD; EAHAD-VWD-LOVD).[18] The patients' phenotypic and molecular data were categorized according to the VWD type and used to determine genotype-phenotype correlations, which were established in 94.6% of subjects (Figure 4). Distribution of the families by VWD subtypes is depicted in Figure 5.

Molecular epidemiology per type of VWD

VWD type 1 was diagnosed in 159 patients. In total, 216 variants were identified in 146 patients (91.8%; *Online Supplementary Table S3*). Eighty-five different variants (54 undescribed in EAHAD-VWD-LOVD)[18] were scattered over *VWF*, being more frequent in the D3 and D4 domains (45.9%). The five most recurrent mutations

Table 1. *VWF* recurrent mutations from 480 patients included in PCM-EVW-ES, classified by VWD types.

Type 1			Type 3			Type 2A			Type 2A/2M			Type 2B			Type 2M			Type 2N		
Mutations	#P	#F	Mutations	#P	#F	Mutations	#P	#F	Mutations	#P	#F	Mutations	#P	#F	Mutations	#P	#F	Mutations	#P	#F
p.Arg1205His	35	15	c.4146G>T p.Asn2066Ser p.Tyr126ThrfsTer49	8	7	p.Arg1374His	30	7	p.Arg1374Cys	22	6	p.Arg1306Trp	11	7	p.Arg1399His	6	6	p.Arg854Gln	15	7
p.Arg924Gln	9	8	p.Gln1311Ter	7	5	p.Ser1506Leu	10	7	p.Arg1315Cys	7	4	p.Arg1308Cys	12	7	p.Gly1415Asp	5	3	p.Arg816Trp	5	4
p.Tyr1584Cys	7	7	p.Pro2063Ser	6	5	p.Ile1628Thr	11	6	p.Arg763Ser	3	2	p.Val1316Met	5	4	p.Val1409Phe	7	2	-		
p.Leu1733Pro	11	5	p.Gln2783Ter	4	3	p.Arg1597Trp	9	5	-			-			-			-		
p.Pro2063Ser	11	5	p.Arg324Ter	4	3	p.Cys1149Arg	11	4	-			-			-			-		
% of type 1 families	39%		% of type 3 families	45%		% of type 2A families	53%		% of type 2A/2M families	92%		% of type 2B families	78%		% of type 2M families	48%		% of type 2N families	100%	

Last lane shows the % of families presenting the most recurrent mutations identified in each type of VWD. #F indicates number of families and #P, number of patients where the corresponding mutation was found.

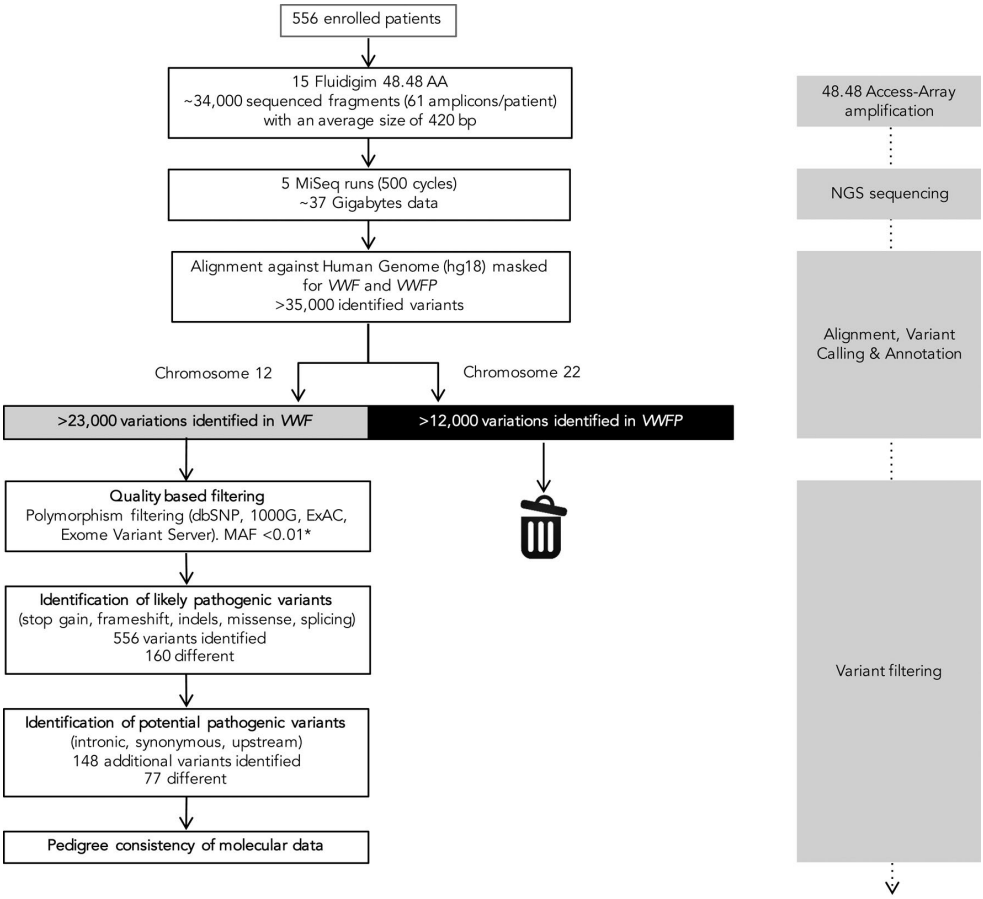
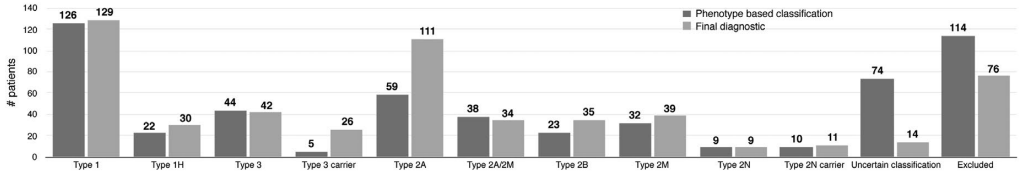


Figure 1. Flowchart depicting the molecular analysis and identification of potential mutations in individuals enrolled in the study. *VWF* from 48 patients were simultaneously amplified in a 48.48 Access-Array. Alignment, variant calling and annotation of each variant identified was performed by the CLC Genomic Workbench software. This analysis allowed selection of mutations/variants aligned against *VWF* (shown in gray) and elimination of those aligned against *VWFP* (shown in black). Variant filtering was performed by the Variant Studio software. *MAF<0.01 for all variants except three (p.Arg854Gln, p.Arg924Gln, p.Pro2063Ser). *VWF*: von Willebrand factor gene; *VWFP*: von Willebrand factor pseudogene; dbSNP: dbSNP database; 1000G: 1000 Genomes Project; ExAC: Exome Aggregation Consortium; MAF: minor allele frequency; NGS: next-generation sequencing.

(Table 1) had been previously reported in type 1 VWD, except for p.Leu1733Pro, with a deleterious *in silico* score of 4. Patients with the p.Arg1205His, p.Pro1824His, or p.Leu1733Pro mutations showed the lowest levels of VWF:Ag and VWF:RCo (mean 9.1 and 7, respectively). In 30 patients subclassified as VWD 1H,[14,19] the most frequent mutation was p.Arg924Gln (six patients), previously described as a mutation and a polymorphism (*Online Supplementary Results*). With the exception of 13 patients with no detected variant (mean VWF:Ag=31%; range 16-48; 78% blood group O) and patient 01030 who carried the classic p.Arg854Gln 2N mutation, a good phenotype-genotype correlation was established in 145 patients.

Forty-two patients classified as type 3 VWD (mean VWF:Ag=1.2%; mean BS=12.6) and 26 as carriers (mean VWF:Ag=55%; mean BS=1.7) were present in 51 different families (*Online Supplementary Table S4*). In total, 60 different potential mutations were identified throughout *VWF*, and 34 had not been previously described in EAHAD-VWD-LOVD,[18] of which 12 were mutations causing null alleles. Genotyping was clear and correlated with the clinical and laboratory data in all patients except three: patients 05011 and 10011 with a mutation (p.Cys1946Phe and c.5455+1G>A, respectively) in heterozygous state, and patient 10006 with a homozygous nonsense mutation (p.Gln2470Ter) and normal laboratory levels (pending new analysis from a freshly obtained sample). Excluding these three cases, the 74.4% of type 3 patients had mutations leading to incontrovertible null alleles in both chromosomes (*Online Supplementary Figure S3A*).

Type 2A VWD was finally diagnosed in 111 patients (detailed classification is disclosed in *Online Supplementary Results*). In total, 158 variants were identified (*Online Supplementary Table S5*) and most were clustered in the A1 and A2 domains (38 and 57 potential mutations, respectively) encoded by exon 28. Ninety-eight patients (88%) showed a mutation classified as 2A in EAHAD-VWD-



Phenotype based classification	126	22	44	5	59	38	23	32	9	10	74	114
Molecular Study												
Out	13	8	3	6	2	7	2	1	0	0	68	43
To	Type 1H Type 3 Carrier Type 2M* Type 2B	Type 1 Type 3 Carrier UC	Excluded	Type 1 Type 1H	Excluded Type 2B	Type 1 Type 2A	Type 2A	Type 2A			Type 1 Type 3† Type 3 Carrier Type 2A Type 2A/2M Type 2B Type 2M Excluded (AVWS)	Type 1 Type 1H Type 3 Carrier Type 2A Type 2M Type 2N Carrier
In	16	16	1	27	54	3	14	8	0	1	8	5
From	Excluded Type 1H Type 3 Carrier Type 2A/2M UC	Excluded Type 1	UC†	Excluded Type 1 Type 1H UC	Excluded Type 2A/2M Type 2B Type 2M UC	UC	Type 1 Type 2A UC	Excluded Type 1* UC		Excluded	Type 1H	Type 3 (AVWS) Type 2A UC (AVWS)
Final diagnostic	129	30	42	26	111	34	35	39	9	11	14	76

Figure 2. Classification of VWD patients based on central phenotypic diagnosis and final assignment according to the genotype-phenotype correlation. The dataset highlights the number of patients reincluded or reclassified after molecular study, from the initial classification based on central phenotypic results (shown in black) to a definite, refined classification based on molecular data analysis together with phenotypic results (shown in gray). *Patients diagnosed initially as type 1 VWD and reclassified to type 2M VWD due to the presence of collagen binding mutations. †Patient with uncertain classification reclassified to type 3 VWD due to the presence of nonsense mutation in homozygous state. This is a discrepant case since laboratory levels do not correlate and it is pending a new analysis from a freshly obtained sample. Out: number of patients removed from this subtype; To: final definite classification; In: number of patients reclassified to this subtype; From: previous phenotypic classification; UC: uncertain classification; AVWS: acquired von Willebrand Syndrome.

LOVD,[18] and the remaining 13 patients had novel candidate mutations in one of the 2A-associated domains (D2, D3, A1, A2, and CK). Hence, an unflawed phenotype-genotype correlation was found in all except four patients: patient 42010, who showed a previously described 2A mutation (p.Arg1527Gln)[20] with healthy compatible VWF:Ag and VWF:RCo levels (pending new analysis of a freshly obtained sample), patient 03012, who had a large in-frame deletion affecting residue 386-1873 (domains D2-A3), although this type of mutation seldom explains a 2A phenotype[21] and additional studies are needed to investigate the molecular origin, and patients 02071 and 02072, who had a nearly normal multimeric pattern and VWF:RCo/VWF:Ag>0.7 due to the heterozygous missense mutations, p.Arg976Cys (described as 2A/IIE)[22] and p.Pro2063Ser.

Thirty-four patients were classified as type 2A/2M VWD (Figure 3 and *Online Supplementary Table S6*). Remarkably, all families save three presented one of the widely described p.Arg1315Cys and p.Arg1374Cys mutations, located in the A1 domain.[23,24] These two controversial mutations were difficult to classify in the light of previous reports and were assigned to 2A/2M, mainly because of their abnormal VWF multimers in medium-resolution gels.[25] The new p.Arg763Ser potential mutation, involved in proteolytic processing of the VWF precursor, was found in heterozygous state in two families with symptomatic patients (BS >4), a smeary multimeric pattern, and VWF:FVIIIB <0.7. Of note, mutations in the same amino acid (p.Arg763Gly) have been described in type 2N, but heterozygous patients for this mutation were asymptomatic and classified as 2N carriers.[26,27] The remaining family showed p.Cys2491Arg, described as type 3, which involves loss of a cysteine essential for VWF secretion, but these patients were classified as 2A/2M based on their multimeric pattern. Taking these data together, a complete genotype-phenotype correlation could be established for all 2A/2M patients.

Thirty-five patients were diagnosed as having type 2B VWD (detailed classification is disclosed in *Online Supplementary Results* and *Online Supplementary Table S7*). Interestingly, among the 16 residues associated with type 2B VWD, one out of five of them—p.Pro1266, p.Arg1306, p.Arg1308, p.Val1316 or p.Pro1337—was found to be affected. A good phenotype-genotype correlation could be established for all patients, as most showed a loss of high-molecular-weight multimers and discordance between VWF:Ag and VWF:RCo levels (mean ratio=0.51; range 0.19-1.1) and a classical type 2B mutation. A normal multimeric pattern and VWF:RCo/VWF:Ag>0.7 were observed in only two families. As expected, the classical p.Pro1266Leu and p.Pro1266Gln molecular defects were found, both being responsible for type 2B Malmö/New York VWD.[28,29]

Thirty-nine patients were classified as type 2M VWD (*Online Supplementary Table S8*). Twenty-one different missense variants were found, ten undescribed in EAHAD-VWD-LOVD.[18] Of interest, in eight patients with a high

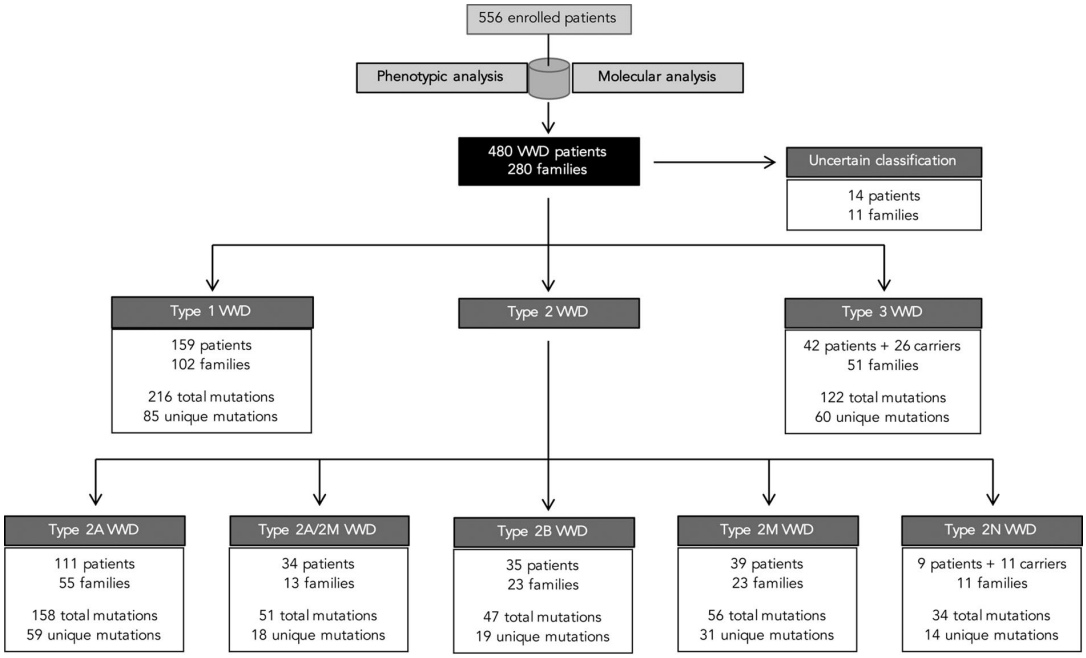


Figure 3. Summary of the final diagnosis of patients in terms VWD type. On the basis of phenotype data, 98 of the 556 recruited individuals did not meet any inclusion criteria and were initially excluded. Following the molecular analysis, 43 of these patients were reincluded due to the presence of a candidate mutation in *VWF*. The remaining 55 patients did not meet any inclusion criteria. Moreover, molecular analysis prompted the exclusion of 21 additional individuals: eight HA patients, five HA carriers, and eight patients finally diagnosed as AVWS and confirmed by the absence of mutations in *VWF* or *GP1BA*. Of note, 280 families were finally included (shown in black) and nine of them had members with different VWD types. In those particular cases, the family may be counted more than once; that is, within each VWD type where a family member was classified. VWD: von Willebrand disease.

VWF:RCo/VWF:Ag ratio, collagen-binding mutations were identified: p.Arg1399His, p.Ser1731Thr, and the novel p.Arg1395Trp.[30] All patients showed mutations in the 2M-associated domains A1 and A3, with the exception of three patients who had mutations in other domains, and patient 27022 (carrying p.Ser1731Thr, a collagen I/III-binding mutation, but with VWF:CB=41), considered to have a discrepant genotype-phenotype correlation.

A perfect phenotype-genotype correlation was established in all type 2N VWD patients (*Online Supplementary Table S9*). Two classical 2N mutations were present in all patients: p.Arg816Trp in exon 19, and p.Arg854Gln in exon 20, the latter seen in 75% of cases. Different combinations of mutation types were found (*Online Supplementary Figure S3B*): five patients homozygous for a 2N mutation, and four patients compound heterozygous in trans for p.Arg854Gln and a nonsense mutation or p.Gln895His, reported to cause low *VWF* expression at the messenger ribonucleic acid (mRNA) level.[31]

In 14 patients the classification was uncertain (*Online Supplementary Table S10*) because phenotyping or genotyping data were insufficient or inconclusive to indubitably establish the classification. This occurred in eight patients with novel variants of unknown significance (VUS) who had a normal or nearly normal phenotype, and six patients with an abnormal multimeric pattern, VWF:RCo/VWF:Ag <0.7, and a novel candidate mutation. In these cases, it was difficult to distinguish between 2A or 2B VWD since RIPA assay was not assessed.

Molecular epidemiology per type of mutation

A total of 237 different variants were identified in our cohort (*Online Supplementary Figure S4A*) and a potential ancestral origin was established in 19 of them. All variants, classified by type, are listed in *Online Supplementary Tables S11-S17*.

In total, 119 different missense candidate mutations were identified, scattered over all domains of *VWF* (33 of 52 exons). Sixty were not included in EAHAD-VWD-LOVD[18] (*Online Supplementary Figure S4B*) and the mean *in silico* score in this group was 3.50 (standard deviation (SD) 1.60). The remaining 59 mutations, previously found in VWD patients, had a similar score of 3.54 (SD 1.62). Twenty-three different nonsense mutations were found, 12 unrecorded in EAHAD-VWD-LOVD.[18] As expected, nonsense mutations predominantly occurred in type 3 VWD (41 patients). However, we also found this type of mutation in nine type 1 VWD patients, two type 2N, and one type 2A. Of particular note, several mutations were detected in different VWD types. For example, p.Arg324Ter was found in types 3 and 2N, and p.Gln840Ter, p.Gln1311Ter and p.Gln2470Ter were found in types 1 and 3. Five new inframe candidate mutations (two insertions and three deletions) causing different types of VWD were identified. Thirteen potential splicing

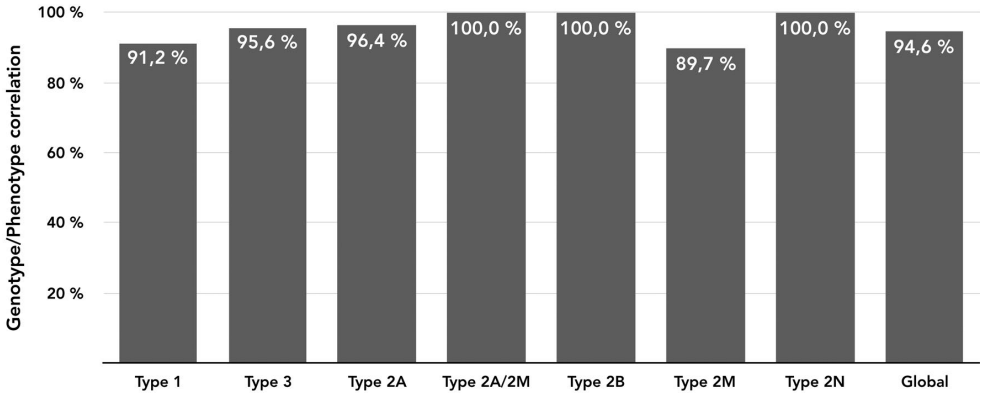


Figure 4. Schematic representation of phenotype/genotype correlations by VWD type. The correlation between genotype and phenotype in each patient was based on the concordance between the results of the phenotypic test panel and the results of the genetic analysis.

mutations were found, eight undescribed in EAHAD-VWD-LOVD.[18] Their deleterious effect was assessed by five *in silico* tools: 11 had a predicted effect on splicing, but two of them, those farthest from the exon to be precise, were predicted to have no significant impact on splicing. Ten different upstream variants were found in 20 families with all types of VWD. Only c.-1896C>T and c.-1873A>G have been previously described,[32,33] and these were found in type 2A/2M and 2M, respectively. According to their genomic localization, all but the three most upstream (c.-3151T>G, c.-2692C>T, c.-2627C>T) were mapped in regions with well-defined or potential regulatory elements. Sixteen synonymous and 51 intronic variants were identified, but only 5 and 3, respectively, were recorded in EAHAD-VWD-LOVD.[18] Although it is difficult to demonstrate, any nucleotide change can be considered a potential candidate to alter splicing, especially when an intronic or synonymous variant appears next to an intron/exon junction.[34,35]

Discussion

This report describes the molecular study of *VWF* in the largest prospective multicenter cohort of VWD patients to date, made possible by the development and successful application of NGS technology. The method presented herein has also been used in a smaller Portuguese cohort, with a similar diagnostic yield.[16] Heretofore, *VWF* genetic analysis in large cohorts has been performed by Sanger sequencing.[32,33,36,37] In the study herein, considerable time and cost savings were attained by combining two high-throughput technologies, nanofluidics and NGS. The hands-on time and instrument run times for sequencing the 556 patients initially enrolled can be measured in days or weeks rather than months, whereas the cost per sample in our setting was less than $70, 10-fold lower than Sanger sequencing.

The method developed for *VWF* sequencing is fast and cost-effective for large patient groups; hence, it would seem inappropriate for routine diagnostic laboratories that accumulate small numbers of patients. Nonetheless, NGS can also be applied to individual samples by using the gene panel approach.[38] In our opinion, the debate about whether genetic analysis is appropriate for all types of VWD[12] will become obsolete in the light of this new technological scenario. It is reasonable to believe that NGS will be progressively incorporated in routine VWD diagnostics.

Centralized comprehensive population studies, although logistically complex, would be the best way to obtain a clear picture of this disease and investigate connections between the multifactorial parameters that influence the diagnosis and etiology.[12] In this study, 704 variants (237 unique) were compiled (*Online Supplementary Tables S11-S17*), whereas the total number of entries in EAHAD-VWD-LOVD[18] is about 1,200 (708 unique), an indication of the effectiveness of this approach for the genetic study of large cohorts. In contrast to other studies, the present project included patients with all VWD types and used an identical protocol (whole *VWF* sequencing) for their genetic analysis. The mutation detection rate was high compared with that of previous studies, being close to 100% in some VWD types (Figure 3). Candidate mutations were identified in 91.8% of type 1 individuals *versus* the 65-82% reported in previous studies,[32,33,36,39] a result that may be related to the limiting recruitment of patients with VWF level <30%.

With regard to the most recurrent mutations identified (Table 1), the Vicenza mutation (p.Arg1205His) proved to be the most common in our type 1 patients, in contrast to previous studies where p.Tyr1584Cys was prevalent.[32,33,40] In type 3 VWD, p.Gln1311Ter[41] was one of the most frequent mutations detected, as was also seen in the Portuguese population study,[16] whereas in patients from northern Europe, p.Arg2535Ter is predominant. As to type 2, the most common mutations found were similar to those reported in previous cohort studies.

Perhaps more important than elucidation of the molecular epidemiology of VWD were the data obtained by NGS, that can resolve many of the drawbacks and limitations of phenotyping. First, molecular characterization of VWD patients enables accurate classification: 27.5% of our patients (153/556) were re-diagnosed after the initial

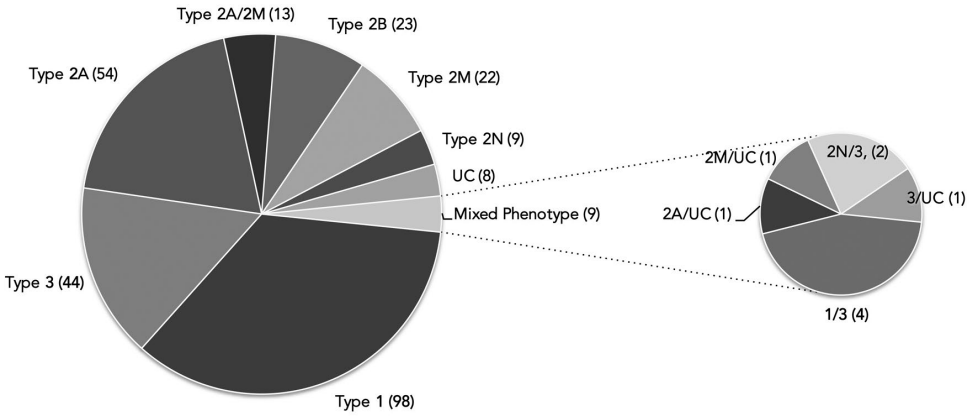


Figure 5. Classification of the 280 families included according to VWD type. Numbers in parentheses refer to the total of families included in each VWD type. Of note, nine of the 280 families included in the PCM-EVW-ES presented a mixed phenotype (families in which some members had different VWD subtypes). UC: patients with an uncertain classification.

phenotypic study, 43 reincluded, five patients excluded, and 105 reclassified. Of note, 68 patients with uncertain classification were definitively classified by NGS, as in the case of 25 type 2A patients and ten type 2B missing previous RIPA data. Furthermore, four patients whose phenotype pointed to VWD type 1 were found to have 2B Malmö or collagen-binding mutations (Figure 2), the latter being undetectable by standard laboratory tests unless specific binding capacity for collagen types III, IV and VI is performed. Accurate classification is particularly relevant for genetic counseling and detection of carriers in VWD types with autosomal recessive inheritance. We identified 11 asymptomatic carriers 2N and 26 type 3 with borderline plasmatic levels, who would likely have been excluded if the diagnosis had been based only on clinical and laboratory data. Knowledge of the molecular defect also led to diagnostic reassessment of some type 1 patients who could be considered carriers of type 3 mutations (e.g., patient 06024 with the p.Arg2434Ter mutation).

Second, genetic study showed that 37.5% of patients in our cohort had more than one variation, a finding essential to unravelling the potential contribution of different variants to the final phenotype. The situation of two 2M patients is paradigmatic in this regard: 44003, with the p.Gly1415Asp mutation and very low FVIII levels (11%) due to a 2N mutation in heterozygous state, and 03024, who had the p.Val1409Phe plus the p.Arg1399His mutation in trans, which explains the reduction in collagen-binding affinity in this patient.

Third, genetic analysis in large cohorts could shed light on differentiating a variant previously described as a mutation or polymorphism, such as p.Pro1162Leu caused by the c.3485C>T substitution, described as a polymorphism in African Americans.[42] Conversely, in our cohort p.Pro1162Leu was caused by c.3485_3486delinsTG (predicted by *in silico* analysis with effect on splicing), leading to type 3 phenotype in homozygous state.[16] These data suggest that different variants resulting in identical amino acid changes may have different consequences at the transcriptional level. Additionally, regarding the controversial p.Pro2063Ser variant, a recent report by Kasatkar *et al*.[43] has provided evidence that p.Pro2063Ser in homozygous state causes type 3 VWD. Hence, five patients in our cohort heterozygous for p.Pro2063Ser were classified as type 3 carriers.

Fourth, genetic studies can offer valuable data regarding the molecular pathophysiological mechanisms responsible for the symptoms observed, such as defects in the structure, intracellular transport, or secretion of VWF. The location of new mutations in certain domains or specific amino acids helps to predict their potential effect, particularly if previous *in vitro* studies are performed. Examples of this are the new mutations p.Arg273Pro, located in the propeptide and related with formation of disulphide-linked multimers,[44] p.Arg763Ser, located in the VWF propeptide cleavage site,[26] and p.Arg1395Trp, which affects an essential amino acid for collagen IV binding.[30]

In addition, although this approach cannot resolve some discrepant cases and those in which no candidate mutation is found (3.1% of our cohort), exome and genome studies by NGS are useful to identify modifying mutations in other genes[45] and reveal structural variations undetectable by MLPA.[46] It is also important to point out that the technique has provided molecular data for more than 100 cSNPs described in *VWF*, which will be further analyzed to determine their potential influence on VWF:Ag, FVIII:C and bleeding in our cohort, as has been done in some healthy populations.[42,47] If some of these SNPs are found to contribute to the disease, they could be redefined as hemorrhagic risk polymorphisms (similar to the thrombotic risk polymorphisms).

The final obstacle to an accurate diagnosis is the uncertain pathogenicity of novel variations. *In silico* analysis is considered a suitable supporting tool for genetic diagnosis, and we analyzed all variants using algorithms recommended in specialized guidelines.[48] However, the predictive capacity of these programs remains modest: some well-established deleterious mutations have a lower overall score (e.g., p.Leu1580Pro, score=2) than some polymorphisms (e.g., p.Cys325Phe, predicted to be deleterious by

SIFT and PolyPhen).[49] Therefore, in order to unequivocally determine the potential deleterious effect of new variants, *in vitro* functional studies remain essential.[50]

Herein, we present the largest prospective study to date of a cohort of VWD including an exhaustive description of their molecular epidemiology. This data will improve knowledge about the molecular mechanisms that contribute to the complexity of disease diagnosis and will help to elucidate the relationship between bleeding parameters and the patients' laboratory and genetic profiles. The approach used implies a change in the diagnostic paradigm of VWD, which is also occurring in other genetic diseases. The cost reduction and simplification of genetic analysis by NGS will likely lead to the use of molecular study as an additional first-line routine diagnostic test, with the ultimate aim of providing more personalized and effective care for this patient population.

Funding

We are indebted to Baxalta US Inc., now a part of Shire, for its support of the PCM-EVW-ES (H13-000845 Grant). This work was also supported by the Spanish Ministerio de Economía y Competitividad (MINECO)-Instituto de Salud Carlos III (ISCIII) (PI12/01494, PI15/01643 and RD12/0042/0053). We thank Technoclone GmbH, Austria, Labclinics, SA, Spain for donating the VWF:CB (type VI collagen) kit. We are very grateful for the kind collaboration of the participating patients and their families. CIBERCV is an initiative of ISCIII co-financed by Fondo Europeo de Desarrollo Regional (FEDER) a way to build Europe.

References

1. Sadler JE, Budde U, Eikenboom JC, et al. Update on the pathophysiology and classification of von Willebrand disease: a report of the Subcommittee on von Willebrand Factor. J Thromb Haemost. 2006;4(10):2103-2114.
2. Rodeghiero F, Castaman G, Dini E. Epidemiological investigation of the prevalence of von Willebrand's disease. Blood. 1987;69(2):454-459.
3. Ruggeri ZM. Von Willebrand factor, platelets and endothelial cell interactions. J Thromb Haemost. 2003;1(7):1335-1342.
4. Vlot AJ, Koppelman SJ, Meijers JC, et al. Kinetics of factor VIII-von Willebrand factor association. Blood. 1996;87(5):1809-1816.
5. Castaman G, Hillarp A, Goodeve A. Laboratory aspects of von Willebrand disease: test repertoire and options for activity assays and genetic analysis. Haemophilia. 2014;20 Suppl 4:65-70.
6. Hamilton A, Ozelo M, Leggo J, et al. Frequency of platelet type versus type 2B von Willebrand disease. An international registry-based study. Thromb Haemost. 2011;105(3):501-508.
7. Batlle J, Pérez-Rodríguez A, Costa-Pinto J, Lourés E, Rodriguez-Trillo A, López-Fernández MF. Diagnosis and management of von Willebrand disease in Spain. Semin Haemost Thromb 2011; 37(5):503-510. Semin Haemost Thromb. 2011;37(5):503-510.
8. Flood VH. New insights into genotype and phenotype of VWD. Hematology Am Soc Hematol Educ Program. 2014;2014(1):531-535.
9. Mancuso DJ, Tuley EA, Westfield LA, et al. Human von Willebrand factor gene and pseudogene: structural analysis and differentiation by polymerase chain reaction. Biochemistry. 1991;30(1):253-269.
10. Corrales I, Ramirez L, Altisent C, Parra R, Vidal F. Rapid molecular diagnosis of von Willebrand disease by direct sequencing. Detection of 12 novel putative mutations in VWF gene. Thromb Haemost. 2009;101(3):570-576.
11. Corrales I, Catarino S, Ayats J, et al. High-throughput molecular diagnosis of von Willebrand disease by next generation sequencing methods. Haematologica. 2012;97(7):1003-1007.
12. Goodeve AC. The genetic basis of von Willebrand disease. Blood Rev. 2010;24(3):123-134.
13. Desai A, Jere A. Next-generation sequencing: ready for the clinics? Clin Genet. 2012;81(6):503-510.
14. Batlle J, Perez-Rodriguez A, Corrales I, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): Proposal for a new diagnostic paradigm. Thromb Haemost. 2016;115(1):40-50.
15. Consortium EP, Bernstein BE, Birney E, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.
16. Fidalgo T, Salvado R, Corrales I, et al. Genotype-phenotype correlation in a cohort of Portuguese patients comprising the entire spectrum of VWD types: impact of NGS. Thromb Haemost. 2016;116(1):17-31.
17. Vidal F, Julia A, Altisent C, Puig L, Gallardo D. Von Willebrand gene tracking by single-tube automated fluorescent analysis of four short tandem repeat polymorphisms. Thromb Haemost. 2005;93(5):976-981.
18. European Association for Haemophilia and Allied Disorders (EAHAD). Coagulation Factor Variant Databases. October 01, 2010 ed.
19. Montgomery RR, Christopherson P, Bellissimo DB, et al. The complete type I VWD cohort of the Zimmerman Program for the molecular and clinical biology of VWD - phenotypic assignment, mutation frequency, and bleeding assessment. Blood. 2013;122(21):332.
20. Ahmad F, Jan R, Kannan M, et al. Characterisation of mutations and molecular studies of type 2 von Willebrand disease. Thromb Haemost. 2013;109(1):39-46.
21. Haberichter SL, Allmann AM, Jozwiak MA, Montgomery RR, Gill JC. Genetic alteration of the D2 domain abolishes von Willebrand factor multimerization and trafficking into storage. J Thromb Haemost. 2009;7(4):641-650.
22. Schneppenheim R, Michiels JJ, Obser T, et al. A cluster of mutations in the D3 domain of von Willebrand factor correlates with a distinct subgroup of von Willebrand disease: type 2A/IIE. Blood. 2010;115(23):4894-4901.
23. Penas N, Perez-Rodriguez A, Torea JH, et al. von Willebrand disease R1374C: type 2A or 2M? A challenge to the revised classification. High frequency in the northwest of Spain (Galicia). Am J Hematol. 2005;80(3):188-196.
24. Batlle J, Perez-Rodriguez A, Franqueira MD, Lopez-Fernandez MF. Type 2M von Willebrand disease: a variant of type 2A? J Thromb Haemost. 2008;6(2):388-390.
25. Gadisseur A, van der Planken M, Schroyens W, Berneman Z, Michiels JJ. Dominant von Willebrand disease type 2M and 2U are variable expressions of one distinct disease entity caused by loss-of-function mutations in the A1 domain of the von Willebrand factor gene. Acta Haematol. 2009;121(2-3):145-153.
26. Hilbert L, Nurden P, Caron C, et al. Type 2N von Willebrand disease due to compound heterozygosity for R854Q and a novel R763G mutation at the cleavage site of von Willebrand factor propeptide. Thromb Haemost. 2006;96(3):290-294.
27. van den Biggelaar M, Meijer AB, Voorberg J, Mertens K. Intracellular cotrafficking of factor VIII and von Willebrand factor type 2N variants to storage organelles. Blood. 2009;113(13):3102-3109.
28. Weiss HJ, Sussman, II. A new von Willebrand variant (type I, New York): increased ristocetin-induced platelet aggregation and plasma von Willebrand factor containing the full range of multimers. Blood. 1986;68(1):149-156.
29. Weiss HJ. Type 2B von Willebrand disease and related disorders of patients with increased ristocetin-induced platelet aggregation: what they tell us about the role of von Willebrand factor in hemostasis. J Thromb Haemost. 2004;2(11):2055-2056.
30. Flood VH, Schlauderaff AC, Haberichter SL, et al. Crucial role for the VWF A1 domain in binding to type IV collagen. Blood. 2015;125(14):2297-2304.
31. Cabrera N, Casana P, Cid AR, Haya S, Moret A, Aznar JA. Novel missense mutation c.2685G>C (p.Q895H) in VWF gene associated with very low levels of VWF mRNA. Ann Hematol. 2009;88(3):245-247.
32. Goodeve AC, Eikenboom J, Castaman G, et al. Phenotype and genotype of a cohort of families historically diagnosed with type 1 von Willebrand disease in the European study, Molecular and Clinical Markers for the Diagnosis and Management of Type 1 von Willebrand Disease (MCMDM-1VWD). Blood. 2007;109(1):112-121.
33. James PD, Notley C, Hegadorn C, et al. The mutational spectrum of type 1 von

Willebrand disease: Results from a Canadian cohort study. Blood. 2007; 109(1):145-154.
34. Yadegari H, Biswas A, Akhter MS, et al. Intron retention resulting from a silent mutation in the VWF gene that structurally influences the 5' splice site. Blood. 2016; 128(17):2144-2152.
35. Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet. 2011;12(10):683-691.
36. Cumming A, Grundy P, Keeney S, et al. An investigation of the von Willebrand factor genotype in UK patients diagnosed to have type 1 von Willebrand disease. Thromb Haemost. 2006;96(5):630-641.
37. Veyradier A, Boisseau P, Fressinaud E, et al. A laboratory phenotype/genotype correlation of 1167 French patients from 670 families with von Willebrand disease: a new epidemiologic picture. Medicine (Baltimore). 2016;95(11):e3038.
38. Borràs N, Corrales I, Ramírez L, Altisent C, Parra R, Vidal F. Diseño, optimitzación y validación de un panel de secuenciación de 23 genes como herramienta de diagnóstico e investigación de las coagulopatías congénitas. XXXI Congreso Nacional de la Sociedad Española de Trombosis y Hemostasia; 2015 October; Valencia, Spain: Thrombosis and Haemostasis; 2015. p. 95.
39. Flood VH, Christopherson PA, Gill JC, et al. Clinical and laboratory variability in a cohort of patients diagnosed with type 1 VWD in the United States. Blood. 2016; 127(20):2481-2488.
40. Bowen DJ, Collins PW, Lester W, et al. The prevalence of the cysteine1584 variant of von Willebrand factor is increased in type 1 von Willebrand disease: co-segregation with increased susceptibility to ADAMTS13 proteolysis but not clinical phenotype. Br J Haematol. 2005; 128(6):830-836.
41. Casana P, Martinez F, Haya S, Lorenzo JI, Espinos C, Aznar JA. Q1311X: a novel nonsense mutation of putative ancient origin in the von Willebrand factor gene. Br J Haematol. 2000;111(2):552-555.
42. Bellissimo DB, Christopherson PA, Flood VH, et al. VWF mutations and new sequence variations identified in healthy controls are more frequent in the African-American population. Blood. 2012; 119(9):2135-2140.
43. Kasatkar P, Ghosh K, Shetty S. A common founder mutation p.P2063S in exon 36 of VWF in 11 unrelated Indian von Willebrand disease (VWD) families. Ann Hematol. 2013;92(8):1147-1148.
44. Allen S, Abuzenadah AM, Hinks J, et al. A novel von Willebrand disease-causing mutation (Arg273Trp) in the von Willebrand factor propeptide that results in defective multimerization and secretion. Blood. 2000;96(2):560-568.
45. Souto JC, Almasy L, Soria JM, et al. Genome-wide linkage analysis of von Willebrand factor plasma levels: results from the GAIT project. Thromb Haemost. 2003;89(3):468-474.
46. Suzuki T, Tsurusaki Y, Nakashima M, et al. Precise detection of chromosomal translocation or inversion breakpoints by whole-genome sequencing. J Hum Genet. 2014; 59(12):649-654.
47. Johnsen JM, Auer PL, Morrison AC, et al. Common and rare von Willebrand factor (VWF) coding variants, VWF levels, and factor VIII levels in African Americans: the NHLBI Exome Sequencing Project. Blood. 2013;122(4):590-597.
48. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-424.
49. Wang QY, Song J, Gibbs RA, Boerwinkle E, Dong JF, Yu FL. Characterizing polymorphisms and allelic diversity of von Willebrand factor gene in the 1000 Genomes. J Thromb Haemost. 2013; 11(2):261-269.
50. Corrales I, Ramirez L, Altisent C, Parra R, Vidal F. The study of the effect of splicing mutations in von Willebrand factor using RNA isolated from patients' platelets and leukocytes. J Thromb Haemost. 2011; 9(4):679-688.

Material Suplementari Article 2

SUPPLEMENTAL DATA

SUPPLEMENTAL METHODS

The VWD classification used in this paper was based in Sadler proposal.[1] Type 1 VWD diagnosis is based VWF:Ag, VWF:RCo or/and VWF:CB <30% at the recruitment moment of the study. The qualitative type 2 VWD was defined as VWF:RCo/VWF:Ag ratio <0.7. A minor modification was the inclusion of subtype 1H (historical) [2] to patients that were diagnosed as type 1 VWD, but at the time of enrollment, their central laboratory findings show a very mild decrease or even a normal VWF plasma level.

Bleeding score and central phenotypic assessment

Bleeding score in each patient was calculated using the ISTH Bleeding Assessment Tool.[3]

The central phenotype characterization was based on laboratory assays including FVIII:C, VWF:Ag, VWF:RCo, VWF collagen binding (VWF:CB) and multimeric pattern. The VWF capacity to bind exogenous FVIII (VWF:FVIIIB) was assessed in patients who presented a FVIII:C/VWF:Ag ratio <1, and in those in whom a type 2N mutation was found. Detailed information is available in a previous report.[4]

Sample collection

Genomic DNA was obtained from peripheral blood samples by using the QIAsymphony DNA Midi Kit on a QIAsymphony SP instrument (Qiagen). DNA was quantified by absorbance at 260 nm with a NanoDrop Lite spectrophotometer (Thermo Scientific), and concentration was adjusted to a range of 25 to 50 ng/μL.

Access Array amplification and sequencing

The Access array was primed in a pre-PCR IFC controller AX (Fluidigm). The chip was loaded with samples and PCR reagents and was transferred to the Fluidigm FC1 Cycler using the Fluidigm AA 48x48 Standard v1 protocol (Supplemental Table 1) for a Fusion PCR. Pooled amplicons from each DNA template were harvested in a post-PCR IFC controller AX (Fluidigm), visualized on agarose gel (Figure S1) and quantified by densitometric analysis using ImageJ software (National Institutes of Health). The final pools of up to 192 samples (4 Access Array fusion libraries) were treated for loading onto a MiSeq Reagent Kit v2 cartridge (Illumina). A paired-end, 500 (2x250)-cycle sequencing run was performed on the MiSeq system, following the Illumina standard protocol.

CLC Bio Data analysis parameters

Primer sequences and low-quality regions in specific patients were trimmed and then aligned against the human genome sequence (hg19), masked for the amplified regions of *VWF* and *VWFP.* Since PCR conditions were not stringent enough to completely avoid *VWFP* amplification, double alignment allowed elimination of reads corresponding to the pseudogene (Figure 1). Read mapping was performed with specific parameter settings (mismatch count, 2; indel count, 3; length fraction, 0.7; similarity fraction, 0.9). The Quality-based Variant detection algorithm from the CLC Genomic Workbench was applied for SNV-calling. Analysis parameters were adjusted to obtain optimal performance and maximize sensitivity and specificity in mutation detection. Final analytical settings were established: minimum coverage=10, minimum variant frequency=25%, minimum variant count=5.

VariantStudio data analysis parameters

The following analytical parameters were used in the VariantStudio software to identify putative mutations: 1) frequency, based on minor allele frequency (MAF) from 1000 Genomes (April 2012 v.3) ≤1% for the Global, African, American, Asian and European populations; 2) location, referring to the Variant Effect Predictor v.2.8 database (intronic mutations located >250 bp into the flanking intronic sequence of each exon were discarded); and 3) variant frequency from our local NGS variant database ≤5% .

Nomenclature used and classification of variant pathogenicity

Amino acid numbering and nomenclature was performed in accordance with international recommendations for the description of sequence variants from the Human Genome Variation Society (HGVS; http://www.HGVS.org), applied to genetic variants in hemostasis.[5, 6] Genetic nomenclature remains controversial in many ways, and it is important to specify the exact meaning of the terms used in each case. We have used the term “mutation” to indicate a disease-causing change and the term “variant” or “variants of unknown significance” (VUS) to indicate changes with an unknown functional effect. The term “polymorphism” was used to indicate both a non-disease-causing change and a change found at a frequency of 1% or higher in any population included in the SNP databases.

The following criteria established the pathogenic nature of the variants found: 1) stop/frameshift variants were considered likely to be disease-causing, 2) co-segregation in the family with VWD phenotype, 3) variant identified as disease-causing in the VWD literature or international databases (VWF EAHAD Coagulation Factor Variant Database compiled in LOVD [EAHAD-VWD-LOVD]),[7] 4) *in silico* evaluation findings, and 5) presence of a second mutant

allele in cases of autosomal recessive inheritance.

In addition to mutations that have a very evident effect (eg, stop, frameshift, acceptor or donor splice site) and missense mutations, whose effects are more unpredictable, we have maintained within the PCM-EVW-ES registry all variants in intronic regions and synonymous changes whose MAF in the general population and in the 1000 Genomes project is below 1%.

SUPPLEMENTAL RESULTS

Of the 556 samples analyzed, 1 failed during access array amplification, and 5 resulted in too few reads in the sequencing process, precluding correct analysis. These 6 samples and 9 additional samples with borderline quality values were successfully amplified in a second access array. Moreover, in patients 30035 and 30037, analyzed in the same access array chip and run, mutations were not detected by the standard method because the percentage of reads corresponding to the mutated allele were below the selected threshold (20%) established in the variant-calling algorithm. In both patients, an identical mutation (p.Arg1779Leu) was discovered in further analysis using an NGS custom gene panel.

Type 1H VWD

Six type 1H and three type 1 VWD patients with borderline VWF:Ag levels (mean 38.6; range 13-72) presented the controversial p.Arg924Gln variant, described as a polymorphism and also reported in type 1 and 2N VWD. However, in type 1 patients, p.Arg924Gln was combined with the novel mutation p.Gly2083Asp in 2 patients (21003 and 21005) or with the synonymous mutation c.3390C>T (36008), whereas in type 1H patients, the mutation was always found alone. This suggests that some clinical signs (BS) are not completely reflected in the laboratory parameters or that plasmatic levels fluctuate with the age of patient. Moreover, although *in silico* analysis did not predict a deleterious effect (global score 1; Supplemental Table 11), RNA studies have revealed activation of a cryptic splice in exon 28 and generation of a premature stop codon.[8] Finally, the p.Arg924Gln was found in a type 3 patient (36007) in compound heterozygous state with the missense mutation p.Gly967Val and the intronic VUS c.6799-27C>T. Although we are unable to determine the haplotype, it is likely that the type 3 phenotype may be due to the resulting null allele carrying p.Arg924Gln in combination with the deleterious effect of p.Gly967Val, predicted by *in silico* analysis (score=5). This is a further example illustrating that certain mutations cannot be linked to only a single VWD type,

especially in quantitative defects. Furthermore, the same mutation may cause different phenotypes depending on whether it is alone or combined with others.

Type 2A and 2B VWD reclassification

With the solely phenotypic data, it is difficult to correctly subclassify VWD type 2 because of the heterogeneity of the phenotype, some intrinsic subjectivity in the interpretation of the multimeric pattern and, specially, when RIPA assay is impracticable (since must be performed with a freshly obtained sample).

Along this line, fifty-nine patients were initially classified as **type 2A VWD** due to central phenotypic results. However, after genetic analysis, 2 patients were excluded because the diagnosis of one AVWS (03006) and the reclassification into type 2B of patient 01065. Otherwise, after genetic analysis, 54 patients were included in this subgroup based on the identified mutation: 45 patients were previously unclassified because it was not possible to establish a definite diagnosis in basis of phenotypic study; 2 patients previously misclassified as type 2B due to local erroneous RIPA assay were reclassified as type 2A after molecular analysis; 6 patients misdiagnosed as type 2A/2M or type 2M VWD due to unclear multimeric pattern were finally reclassified as type 2A VWD as all of them presented the previously described p.Arg1374His and; one patient is the discrepant 42010, previously excluded because the laboratory levels were above 30%, but the genetic analysis and the absence of HMWM confirmed a VWD type 2A.

Twenty-three patients were diagnosed as **type 2B VWD** based on the central phenotypic assessment. After genetic analysis, 2 patients were excluded of this subgroup due to local erroneous RIPA assay and the identification of the type 2A mutation p.Arg1374His. Among the 14 patients included in type 2B VWD after genetic analysis, RIPA had been only assayed in one of them (01065) with an observed aggregation at >0.8 mg/mL but not at 0.4 mg/mL, leading to a type 2A classification. Nevertheless, this patient presented the widely described p.Arg1308Cys 2B mutation. A similar weird effect on the aggregation was also previously reported in a family with the type 2B mutation p.Val1316Met.[9] Among the remaining 13 patients 11 presented a lack of the HMWM and were previously unclassified because it was not possible to distinguish, without the molecular data, between 2A and 2B subtypes. Finally, 2 patients with normal multimeric pattern were initially misclassified as type 1 VWD presented the type 2B Malmö/New York mutation (p.Pro1266Gln) and thus were reclassified into this subtype.

These results support the importance of the genetic study as a powerful tool to orient physician in the establishment of a definite sub-classification of type 2 patients.

SUPPLEMENTAL TABLES

Table S1. Thermal cycling protocol (Fluidigim 'AA 48x48 Standard v1') for fusion PCR protocol

PCR Stages		Number of Cycles
50°C	2 minutes	1
70°C	20 minutes	1
95°C	10 minutes	1
95°C 60°C 72°C	15 seconds 30 seconds 1 minute	10
95°C 80°C 60°C 72°C	15 seconds 30 seconds 30 seconds 1 minute	2
95°C 60°C 72°C	15 seconds 30 seconds 1 minute	8
95°C 80°C 60°C 72°C	15 seconds 30 seconds 30 seconds 1 minute	2
95°C 60°C 72°C	15 seconds 30 seconds 1 minute	8
95°C 80°C 60°C 72°C	15 seconds 30 seconds 30 seconds 1 minute	5

Table S2. Sinopsis of outputs per MiSeq Run

Fluidigm AA	Run MiSeq	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Total samples	Samples from registry	First Sequencing Samples
FL1	**Rxn1**	966 +/- 20	84.67 +/- 2.83	0.108 / 0.129	18.05	15.29	78.92	48	48	48
FL2	**Rxn2**	1,248 +/- 58	76.64 +/- 3.46	0.137 / 0.122	20.57	15.79	72.32	96	85	84
FL3										
FL4	**Rxn3**	893 +/- 44	84.02 +/- 4.20	0.116 / 0.077	15.39	12.96	74.48	192	192	191
FL5										
FL6										
FL7										
FL8	**Rxn4**	1,047 +/- 32	75.75 +/- 7.96	0.066 / 0.089	19.77	14.96	72.64	192	189	189
FL9										
FL10										
FL11										
FL12	**Rxn5**	1,018 +/- 43	82.30 +/- 4.21	0.085 / 0.072	18.75	15.45	61.07	192	57	44
FL13										
FL14										
FL15										
Total		**1034.4**	**80.67**		**18.51**	**14.89**	**71.89**	**720**	**571**	**556**

AA indicates Access Array; PF, passing filter; FL, fluidigm; and Rxn, run number.

Table S3. Genotype-phenotype correlation in type 1 and 1H VWD patients

	Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
1	0103	01006	c.7730-56C>T			45	intronic	het	39	F	O+	11	27	32	72	33
	0106	01010	c.4751A>G	**p.Tyr1584Cys**	28		A2	het	55	M	O+	5	27	30	51	32
	0111	01019	c.3223-7_3236dup	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	25		D3	het	52	F	O+	12	17	21	49	20
		01021	c.3223-7_3236dup	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	25		D3	het	11	F	O+	2	18	19	36	19
		01050	c.3223-7_3236dup	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	25		D3	het	9	M	O+	8	29	27	56	24
		01051	c.3223-7_3236dup	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	25		D3	het	36	M	O+	7	31	24	57	25
	0113	01025*	ND						54	M	AB-	5	27	36	45	25
	0114	01028	c.4751A>G	**p.Tyr1584Cys**	28		A2	het	8	M	O+	7	21	33	34	30
			c.5312-19A>C			30	intronic	het								
	0115	01030*	c.2561G>A	p.Arg854Gln	20		D'	het	12	F	O+	2	23	33	51	33
	0119	01037	c.3426T>C	c.3426T>C (p.=)	26		D3	het	16	M	O-	14	22	32	28	22
			c.3485_3486delinsTG	p.Pro1162Leu	26		D3	het								
	0201	02001	c.3614G>A	**p.Arg1205His**	27		D3	het	63	F	A+	14	9.7	2.6	19	6
		02014	c.3614G>A	**p.Arg1205His**	27		D3	het	72	M	A+	12	7.1	5	19	7.3
		02015	c.3614G>A	**p.Arg1205His**	27		D3	het	26	F	A+	8	4.9	8.4	31	4.5
		02016	c.3614G>A	**p.Arg1205His**	27		D3	het	67	M	A+	5	16	9	34	16
		02017	c.3614G>A	**p.Arg1205His**	27		D3	het	56	F	B+	12	5.9	6.5	20	5.4
		02018	c.3614G>A	**p.Arg1205His**	27		D3	het	24	F	O+	8	2.3	0.8	12	2.8
			c.8254-5T>G			51	intronic	het								
		02019	c.2586G>T	c.2586G>T (p.=)	20		D'	het	21	M	AB-	3	6.6	7.3	18	6.2
			c.3614G>A	**p.Arg1205His**	27		D3	het								
		02053	c.3614G>A	**p.Arg1205His**	27		D3	het	58	F	O+	9	10	7.2	18	8.8
		02054	c.3614G>A	**p.Arg1205His**	27		D3	het	26	F	AB-	7	11	8.1	20	10
	0206	02023	c.3614G>A	**p.Arg1205His**	27		D3	het	28	F	O+	18	5.5	8	11	4.2
		02055	c.3614G>A	**p.Arg1205His**	27		D3	het	51	F	B+	14	7.2	6.4	17	7.5
		02056	c.3614G>A	**p.Arg1205His**	27		D3	het	24	F	B-	8	8.8	6.6	18	7.3
		02057	c.3614G>A	**p.Arg1205His**	27		D3	het	53	M	B+	6	6.7	5	11	5.7
	0215	02042	c.1109G>A	p.Cys370Tyr	9		D1	het	41	M	O+	1	25	21	77	27
	0218	02047	c.5471C>A	**p.Pro1824His**	32		A3	het	71	M	O+	9	9.3	6.4	30	6.1
		02048	c.5471C>A	**p.Pro1824His**	32		A3	het	45	F	B+	14	8.3	6.5	39	6.1
		02049	c.5471C>A	**p.Pro1824His**	32		A3	het	43	M	A+	4	11	6.6	37	8
		02050	c.5471C>A	**p.Pro1824His**	32		A3	het	16	M	A+	7	10	6.9	37	7.1
	0220	02052	c.1625C>G	p.Ala542Gly	14		D2	het	25	M	A+	11	15	12	20	16
			c.6187C>T	p.Pro2063Ser	36		D4	het								
	0221	02058	c.3614G>A	**p.Arg1205His**	27		D3	het	64	F	A+	19	7.3	5.9	19	6.3
	0222	02059	c.1946-16_1946-15insCTC			15	intronic	het	38	M	A+	12	5.4	6.1	11	4.5
			c.3614G>A	**p.Arg1205His**	27		D3	het								
		02060	c.3614G>A	**p.Arg1205His**	27		D3	het	67	F	A+	9	8.1	6.1	15	7.2
			c.7438-169G>A			43	intronic	het								
	0223	02061*	ND						29	M	O+	7	18	15.7	41	16
		02062*	ND						24	M	O+	3	16	14.7	49	17

	Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
1	0224	02064	c.3614G>A	**p.Arg1205His**	27		D3	het	32	M	O-	13	4.7	5	16	6
		02065	c.3614G>A	**p.Arg1205His**	27		D3	het	40	M	A-	5	7.2	5.6	24	9.5
		02066	c.3614G>A	**p.Arg1205His**	27		D3	het	66	M	A+	1	7.5	6.1	30	9.7
		02067	c.3614G>A	**p.Arg1205His**	27		D3	het	68	M	A-	4	7.5	5.6	29	10.3
	0225	02068	c.6352C>T c.8155+3G>C	p.Arg2118Trp	37	 50	D4 intronic	het het	58	F	O+	13	11	9.2	46	14
	0226	02069	c.3467C>T c.-3151T>G	**p.Thr1156Met**	26		D3 upstream	het het	45	F	O+	5	30	27	81	40
		02070	c.3467C>T c.-3151T>G	**p.Thr1156Met**	26		D3 upstream	het het	18	F	A+	5	33	29.6	79	44
		02073	c.3467C>T c.-3151T>G	**p.Thr1156Met**	26		D3 upstream	het het	44	F	O+	7	22	16.8	53	26
	0228	02074	c.3426T>C c.6254G>T c.3485_3486delinsTG	c.3426T>C (p.=) p.Cys2085Phe p.Pro1162Leu	26 36 26		D3 D4 D3	het het het	37	F	A+	3	24	21	111	29
	03150	03035	c.3467C>T c.-3151T>G	**p.Thr1156Met**	26		D3 upstream	het het	51	M	ND	4	34	26	43	33
	03230	03044	c.6699_6702dup	p.Cys2235ArgfsTer8	38		null allele	het	27	F	ND	3	22	18	29	17
	0399	03015*	ND						37	F	O+	0	17	17.7	15	16
	0507	05008	c.3614G>A	**p.Arg1205His**	27		D3	het	30	M	A+	2	5.8	5	7.4	6.1
	0511	05012	c.2119T>C	p.Cys707Arg	16		D2	het	10	F	O+	4	15	12.3	45	25
	0602	06002	c.2685G>C c.4751A>G	p.Gln895His **p.Tyr1584Cys**	20 28		D3 A2	het het	39	F	O+	8	25	19.1	88	24
		06003	c.2685G>C c.5665-36T>C	p.Gln895His	20	 33	D3 intronic	het het	74	M	O+	17	31	30	100	33
	0606	06007	c.1156+2T>C			10	intronic	het	54	M	O+	5	28	26.1	56	23
		06008	c.1156+2T>C			10	intronic	het	20	F	A-	5	24	27	54	24
	0607	06009	c.2435del	**p.Pro812ArgfsTer31**	18		null allele	het	60	F	O+	2	23	25	83	23
	0609	06012	c.3363G>T	p.Arg1121Ser	25		D3	het	10	M	O+	10	13	14	40	13
		06013	c.3363G>T	p.Arg1121Ser	25		D3	het	40	M	A+	0	23	23.1	57	23
		06017	c.3363G>T	p.Arg1121Ser	25		D3	het	4	M	A-	8	7.4	6.5	27	9.1
	0617	06024	c.7300C>T	p.Arg2434Ter	43		null allele	het	36	F	O+	1	35	25	69	27
	0703	07003	c.5170+10C>T c.8254-5T>G			29 51	intronic intronic	het het	29	F	O+	5	25	24.9	50	28
	1002	10002	c.3614G>A c.-2077A>T	**p.Arg1205His**	27		D3 upstream	het het	71	M	O+	1	7.8	7.1	8.1	6.4
		10003	c.3614G>A	**p.Arg1205His**	27		D3	het	33	M	A+	1	7.4	7.2	11	6.3
	1011	10013	c.5198T>C	p.Leu1733Pro	30		A3	het	23	F	O+	3	7.2	5.6	19	6.5
	1301	13001	c.7985A>G	p.Lys2662Arg	48		C6	hom	61	F	O+	27	11	9.7	36	12
	1304	13006	c.3614G>A	**p.Arg1205His**	27		D3	het	66	M	A-	8	10	6	16	6.1
	1308	13010	c.1472G>C c.546G>A	p.Arg491Pro c.546G>A (p.=)	13 6		D2 D1	het het	37	F	O+	18	16	15.1	45	14
		13011	c.1472G>C c.7730-177G>T	p.Arg491Pro	13	 45	D2 intronic	het het	39	M	ND	4	25	28.7	81	27

	Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
1	1309	13012	c.3614G>A c.533-43A>C	**p.Arg1205His**	27 5		D3 intronic	het het	42	M	A+	7	14	9	20	11
	1311	13014	c.4751A>G	**p.Tyr1584Cys**	28		A2	het	20	M	O+	6	24	22.5	40	20
	1314	13018	c.4238C>T	**p.Pro1413Leu**	28		A1	het	5	M	ND	0	18	21	39	15
	1402	14002*	ND						32	M	A+	1	28	23	71	25
	1403	14003	c.221-70G>T c.478G>A	p.Gly160Arg	3 5		intronic D1	het het	40	F	O+	6	28	20	61	18
	1616	16001	c.1110-26T>A c.5695T>C	p.Cys1899Arg	9 34		intronic D4	het het	50	F	A+	7	21	16	108	20
		16002	c.1110-26T>A c.5695T>C	p.Cys1899Arg	9 34		intronic D4	het het	27	F	A+	2	22	16	96	22
	1801	18001	c.3614G>A c.6989T>G	**p.Arg1205His** p.Val2330Gly	27 41		D3 C1	het het	50	M	B+	7	10	6.1	17	8.2
		18002	c.3614G>A c.6989T>G	**p.Arg1205His** p.Val2330Gly	27 41		D3 C1	het het	17	M	AB+	0	12	7	16	9.5
		18003	c.3614G>A c.6989T>G	**p.Arg1205His** p.Val2330Gly	27 41		D3 C1	het het	15	M	O+	5	8.6	6.5	12	6.1
	1903	19005*	ND						44	F	O-	8	48	35	50	28
	2103	21003	c.2771G>A c.6248G>A	**p.Arg924Gln** p.Gly2083Asp	21 36		D3 D4	het het	74	M	A+	10	15	14.2	46	13
		21004	c.55+69A>G c.6248G>A	p.Gly2083Asp	2 36		intronic D4	het het	66	F	O+	13	23	19	56	20
		21005	c.2771G>A c.6248G>A	**p.Arg924Gln** p.Gly2083Asp	21 36		D3 D4	het het	63	M	A+	5	13	13	37	12
		21006	c.6248G>A	p.Gly2083Asp	36		D4	het	40	M	O+	8	11	10	29	10.3
	2111	21011	c.246_247insT	p.Ser83Ter	4		null allele	het	54	M	A+	5	12	20.3	33	16
	2120	21023	c.1110-26T>A c.2103C>T c.6248G>A	c.2103C>T (p.=) p.Gly2083Asp	9 16 36		intronic D2 D4	het het het	48	F	O+	2	7.6	6	11	3.9
	2201	22001	c.5170+10C>T c.7730-1G>C		29 45		intronic intronic	het het	41	F	O+	14	18	17.5	35	17
	2206	22006	c.1001G>A c.2561G>A	p.Gly334Glu p.Arg854Gln	9 20		D1 D'	het het	58	M	O+	7	27	20	12	22
		22007*	ND						11	M	O+	1	42	36	53	29
	2207	22011	c.6799-47G>A c.7464C>T c.998-46C>T	c.7464C>T (p.=)	38 44 8		intronic C3 intronic	hom hom hom	67	M	A-	7	18	16	46	14
	2308	23011	c.3614G>A c.-2077A>T	**p.Arg1205His**	27		D3 upstream	het het	64	M	O+	10	7.5	6.6	10	6.9
		23012	c.3614G>A	**p.Arg1205His**	27		D3	het	32	F	O+	12	7	5.3	6.5	6
	2309	23013	c.6911G>A	**p.Cys2304Tyr**	40		C1	het	46	F	O+	0	33	27.5	56	26
	2310	23015	c.2546+132G>A c.3614G>A	**p.Arg1205His**	19 27		intronic D3	het het	74	M	A+	3	18	17.3	19	13
	2703	27030	**c.1533+15G>A**			13	intronic	het	23	M	O+	2	26	28.6	81	30

	Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
1	2707	27012	c.3467C>T c.-3151T>G	**p.Thr1156Met**	26		D3 upstream	het het	24	F	O+	0	33	29	91	32
	2709	27018*	ND						31	M	O+	2	27	28.5	59	25
	2713	27024	c.3467C>T c.-3151T>G	**p.Thr1156Met**	26		D3 upstream	hom hom	56	F	A+	3	5.3	6	38	4.2
	2716	27032	c.2518G>T	p.Glu840Ter	19		null allele	het	49	F	O+	8	31	22.7	72	38
	2902	29002	c.3223-7_3236dup c.5277C>T	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro c.5277C>T (p.=)	25 30		D3 A3	het het	38	F	O+	7	22	17	54	17
	3005	30006	c.5336G>T	p.Arg1779Leu	31		A3	het	70	F	A-	17	24	29.4	30	28
	3006	30008	c.6536C>T	**p.Ser2179Phe**	37		D4	het	24	F	A+	12	18	19.3	35	21
		30009	c.6536C>T	**p.Ser2179Phe**	37		D4	het	8	M	O+	2	9.4	8	12	9.3
	3012	30035	c.5336G>T	p.Arg1779Leu †	31		A3	het	55	F	A+	1	25	16	38	19
	3013	30037	c.5336G>T	p.Arg1779Leu †	31		A3	het	69	M	O+	1	20	17	25	13
	3202	32002	c.7408C>T	p.Gln2470Ter	43		null allele	het	49	F	A+	3	34	30	53	29
	3204	32004	c.2821-123A>C			21	intronic	het	65	F	A-	19	9.8	4	6.3	5
	3205	32005	c.2546+132G>A c.4751A>G	 **p.Tyr1584Cys**	 28	19	intronic A2	het het	28	F	O-	4	30	25	25	32
	3210	32010	c.3931C>T c.55G>A	p.Gln1311Ter **p.Gly19Arg**	28 2		null allele SigNDl peptide	het het	21	M	ND	12	7.2	5.7	21	6.8
	3501	35001	c.5198T>C c.6187C>T c.7771-87G>A	p.Leu1733Pro p.Pro2063Ser	30 36	 46	A3 D4 intronic	het het het	48	F	O+	1	6.6	8	31	8.1
		35004	c.5198T>C c.6187C>T	p.Leu1733Pro p.Pro2063Ser	30 36		A3 D4	het het	16	M	A+	6	5.6	7.9	19	7.8
		35005	c.5198T>C c.6187C>T	p.Leu1733Pro p.Pro2063Ser	30 36		A3 D4	het het	13	F	A+	5	9.6	8.3	25	10
		35006	c.5198T>C c.6187C>T	p.Leu1733Pro p.Pro2063Ser	30 36		A3 D4	het het	19	F	A+	2	9	7.9	21	9.4
	3502	35002	c.3614G>A c.7730-177G>T	**p.Arg1205His**	27 	 45	D3 intronic	het het	51	M	A+	8	7	6	15	10
		35003	c.3614G>A	**p.Arg1205His**	27		D3	het	75	F	A-	2	6.6	8.5	13	11
	3503	35008	c.1625C>G c.5198T>C c.6187C>T	p.Ala542Gly p.Leu1733Pro p.Pro2063Ser	14 30 36		D2 A3 D4	het het het	64	M	A+	2	5.3	5	10	4.9
	3505	35011	c.5198T>C c.6187C>T	p.Leu1733Pro p.Pro2063Ser	30 36		A3 D4	het het	46	M	O+	2	18	14	42	12
		35012	c.5198T>C c.6187C>T	p.Leu1733Pro p.Pro2063Ser	30 36		A3 D4	het het	13	F	O+	1	25	13	52	20
		35013	c.1293+109T>C c.5198T>C c.6187C>T	 p.Leu1733Pro p.Pro2063Ser	 30 36	11	intronic A3 D4	het het het	9	M	O+	1	12	10	25	10
	3601	36001	c.5170+10C>T c.7730-1G>C			29 45	intronic intronic	het het	5	M	A+	2	33	26.4	88	58

	Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
1	3603	36004	c.2546+55G>T			19	intronic	het	69	F	O-	3	22	18.5	19	19
	3606	36008	c.2771G>A	**p.Arg924Gln**	21		D3	het	34	F	O+	8	34	14	37	25
			c.3390C>T	c.3390C>T (p.=)	26		D3	het								
	3702	37002	c.324-2_326dup	p.Ser110GlufsTer13	5		null allele	het	42	F	O+	3	36	35	49	27
	3705	37005	c.2878C>T	p.Arg960Trp	22		D3	het	39	F	O+	9	26	18	44	18
	3706	37006	c.3426T>C	c.3426T>C (p.=)	26		D3	het	45	F	B+	5	7.6	5	8.9	6.9
			c.3614G>A	**p.Arg1205His**	27		D3	het								
			c.7771-49G>A			46	intronic	het								
			c.3485_3486delinsTG	p.Pro1162Leu	26		D3	het								
	3805	38010	c.2878C>T	p.Arg960Trp	22		D3	het	14	F	O+	11	43	27.5	86	36
	3709	37010*	ND						52	F	O+	14	37	27.4	46	27
	3908	39017	c.3390C>T	c.3390C>T (p.=)	26		D3	het	23	F	A+	4	26	17	44	19
		39019	c.3390C>T	c.3390C>T (p.=)	26		D3	het	19	F	O+	4	23	12	30	16
	4001	40001	c.3363G>T	p.Arg1121Ser	25		D3	het	39	F	ND	10	28	14.2	37	21
	4002	40002	c.5198T>C	p.Leu1733Pro	30		A3	het	39	M	O+	9	9.2	6.7	26	7.4
			c.6187C>T	p.Pro2063Ser	36		D4	het								
		40003	c.5198T>C	p.Leu1733Pro	30		A3	het	10	M	ND	8	17	8.9	28	13
			c.6187C>T	p.Pro2063Ser	36		D4	het								
	4501	45001	c.3614G>A	**p.Arg1205His**	27		D3	het	92	F	ND	9	11	6.6	16	9.4
1H	0117	01034	c.3379+1G>A			25	intronic	het	32	M	O+	2	36	49	101	46
			c.-2627C>T				upstream	het								
	0133	01064	c.7082-13G>C			41	intronic	het	23	F	B+	0	47	41	66	37
		01068	c.7082-13G>C			41	intronic	het	16	F	O+	2	31	31	43	26
	0223	02063	c.2771G>A	**p.Arg924Gln**	21		D3	het	56	M	O+	1	35	36	71	38
	03186	03031	c.1625C>G	p.Ala542Gly	14		D2	het	29	F	O+	0	49	38	56	33
			c.533-2A>G			5	intronic	het								
	0616	06023	c.2771G>A	**p.Arg924Gln**	21		D3	het	27	F	O+	-2	44	40	53	33
	0902	09002	c.2771G>A	**p.Arg924Gln**	21		D3	het	30	M	A+	1	104	80	65	77
			c.5312-138G>A			30	intronic	het								
	1204	12008	c.2771G>A	**p.Arg924Gln**	21		D3	het	14	F	O+	2	72	72	80	65
	1211	12018	c.4747C>T	**p.Arg1583Trp**	28		A2	het	73	F	ND	0	63	49	51	42
			c.7730-238G>A			45	intronic	het								
	1702	17004*	ND						52	M	A+	0	46	39	57	30
	1904	19006	c.7408C>T	p.Gln2470Ter	43		null allele	het	41	F	O+	3	46	38	74	31
	2109	21009*	ND						41	F	O+	3	31	40	58	44
	2112	21014	c.1446C>G	p.Ile482Met	13		D2	het	22	F	A+	6	38	40	43	44
	2117	21019	c.2771G>A	**p.Arg924Gln**	21		D3	het	45	M	O+	10	57	50	59	39
	2703	27004	**c.1533+15G>A**			13	intronic	het	25	M	O-	1	37	36	46	36
		27005	**c.1533+15G>A**			13	intronic	het	55	F	O-	-2	47	60.8	90	55
		27006	**c.1533+15G>A**			13	intronic	het	18	F	O+	2	46	44	175	47
		27026	**c.1533+15G>A**			13	intronic	het	59	M	B-	1	37	33	104	43
		27027	**c.1533+15G>A**			13	intronic	het	53	M	A-	2	50	64	129	59
	2709	27017*	ND						31	M	O+	3	33	38,5	66	31

	Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
1H	2717	27035	c.7888-65C>A			47	intronic	het	29	M	O+	7	32	46	73	45
	2718	27036*	ND						55	F	O+	2	41	36	40	30
	2903	29003	c.5170+10C>T			29	intronic	het	29	M	O+	4	58	51	48	56
			c.5278G>A	**p.Val1760Ile**	30		A3	het								
		29004	c.5278G>A	**p.Val1760Ile**	30		A3	het	30	F	B+	7	59	48	42	41
	3207	32007	c.1446C>G	p.Ile482Met	13		D2	het	29	F	O+	4	69	35	46	34
	3607	36009	c.4751A>G	**p.Tyr1584Cys**	28		A2	het	30	F	O+	2	56	48	47	46
			c.5456-62A>G			31	intronic	het								
	3908	39018	c.3390C>T	c.3390C>T (p.=)	26		D3	het	47	F	A+	7	53	43	69	35
	4202	42002	c.2771G>A	**p.Arg924Gln**	21		D3	het	31	F	O	11	186	140	88	130
			c.3390C>T	c.3390C>T (p.=)	26		D3	het								
			c.7771-86G>A			46	intronic	het								
	4203	42003	c.4751A>G	**p.Tyr1584Cys**	28		A2	het	37	F	O-	3	60	48	50	41
	4205	42005	c.3692A>G	p.Asn1231Ser	28		D3	het	38	M	O-	3	48	40.7	36	31

In bold, mutations previously described in type 1 VWD. NI indicates not identified; and ND, not determined.* Patients with a discrepant phenotype-genotype correlation. † Mutation non detected by Fluidigim method due to an imbalance of alleles reads and detected a posteriori by a NGS custom gene panel.

Table S4. Genotype-phenotype correlation in type 3 patients

Family	Patient	Phenotype	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
0121	01040	3	c.970C>T	**p.Arg324Ter**	8		null allele	hom	34	M	A+	26	0	4	3.9	0
0122	01041	3	c.970C>T	**p.Arg324Ter**	8		null allele	hom	48	M	B-	16	0	4	5.4	0
0124	01059	3 (carrier)	c.6187C>T	**p.Pro2063Ser**	36		D4	het	34	M	AB	0	53	45	75	44
			c.970C>T	**p.Arg324Ter**	8		null allele	het								
	01060	3 (carrier)	c.970C>T	**p.Arg324Ter**	8		null allele	het	59	F	O+	2	143	114	100	98
0126	01049	3	c.5170+10C>T			29	intronic	hom	27	F	A+	19	4	1	3.9	0
			c.7730-1G>C			45	intronic	hom								
0202	02004	3 (carrier)	c.2685G>C	p.Gln895His	20		D3	het	73	M	O+	8	69	67.6	146	71
	02007	3 (carrier)	c.2685G>C	p.Gln895His	20		D3	het	8	M	A+	4	53	54.8	115	41
	02009	3 (carrier)	c.2685G>C	p.Gln895His	20		D3	het	7	M	O-	1	61	71	111	46
0204	02020	3	c.3931C>T	**p.Gln1311Ter**	28		null allele	hom	23	M	AB-	24	0	4	2.6	0
0215	02039	3	c.1109G>A	p.Cys370Tyr	9		D1	hom	49	F	O+	23	1.2	6	11	0
	02040	3 (carrier)	c.1109G>A	p.Cys370Tyr	9		D1	het	23	F	O+	0	41	43	104	41
	02041	3 (carrier)	c.1109G>A	p.Cys370Tyr	9		D1	het	20	F	O+	0	35	41	93	33
	02043	3 (carrier)	c.1109G>A	p.Cys370Tyr	9		D1	het	85	F	O+	-2	36	36.5	103	37
	02044	3 (carrier)	c.1109G>A	p.Cys370Tyr	9		D1	het	78	M	O+	-1	40	39	105	42
0217	02046	3	c.4975C>T	**p.Arg1659Ter**	28		null allele	hom	47	F	A-	18	0	4	4.3	0
0219	02051	3	c.3931C>T	**p.Gln1311Ter**	28		null allele	hom	35	F	B+	21	0	4	3.7	0
0301	03018	3	c.546G>A	c.546G>A (p.=)	6		D1	het	20	F	O+	14	0	5	2.9	0
			c.7082-2A>G			41	intronic	het								
			c.8155+3G>C			50	intronic	het								
0302	03016	3	c.5170+10C>T			29	intronic	hom	48	F	A+	13	1.6	5	4.5	0
			c.7730-1G>C			45	intronic	hom								
03127	03037	3	c.3931C>T	**p.Gln1311Ter**	28		null allele	hom	5	M	A	10	0	5	1.9	0
	03038	3 (carrier)	c.3931C>T	**p.Gln1311Ter**	28		null allele	het	32	M	ND	0	67	37	82	56
	03039	3 (carrier)	c.3931C>T	**p.Gln1311Ter**	28		null allele	het	29	F	ND	1	66	34	68	47
0314	03041	3	c.3931C>T	**p.Gln1311Ter**	28		null allele	hom	25	M	O+	7	0	5	2.3	0
03150	03034	3	c.3467C>T	p.Thr1156Met	26		D3	het	6	M	A+	4	3	5	12	2
			c.5170+10C>T			29	intronic	het								
			c.7730-1G>C			45	intronic	het								
			c.-3151T>G				upsteram	het								
	03036	3 (carrier)	c.5170+10C>T			29	intronic	het	40	M	A+	-1	83	65	103	61
			c.7730-1G>C			45	intronic	het								
03167	03032	3	c.2540dupA	**p.Asn847LysfsTer18**	19		null allele	hom	31	M	A+	8	0	4	2	0
	03033	3 (carrier)	c.2540dup	**p.Asn847LysfsTer18**	19		null allele	het	35	M	A+	0	37	30.7	61	29
			c.7771-29G>A			46	intronic	het								
0318	03001	3	c.3835_4105conNG_001212.3:g.6566 6836	p.[(V1279I;Q1311*;I1343V;V1360A;F1369I)]	28		A1	hom	27	F	AB+	8	1	1	4	ND
0321	03046	3	c.100C>T	**p.Arg34Ter**	3		null allele	hom	29	F	O+	12	0	5	3	0
0325	03021	3	c.449T>C	**p.Leu150Pro**	5		D1	het	15	M	A	10	3.7	5	6.7	1.9
			c.7082-2A>G			41	intronic	het								
			c.7730-177G>T			45	intronic	het								
	03022	3 (carrier)	**c.7082-2A>G**			41	intronic	het	48	M	ND	0	88	46	85	48
			c.7730-177G>T			45	intronic	het								
0391	03002	3	c.1625C>G	p.Ala542Gly	14		D2	hom	43	F	A+	8	0	4	5.7	0
			c.533-2A>G			5	intronic	hom								

Family	Patient	Phenotype	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
0501	05001	3	c.1533+1G>A			13	intronic	het	34	M	A+	9	0	5	1.5	0
			c.6256+2dup			36	intronic	het								
0506	05007	3	c.2561G>A	p.Arg854Gln	20		D'	het	14	F	O+	20	2.9	5	5.8	2.7
			c.3319T>C	p.Tyr1107His	25		D3	het								
			c.4146G>T	c.4146G>T (p.=)	28		A1	het								
			c.6197A>G	p.Asn2066Ser	36		D4	het								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	het								
0508	05009	3	c.1992dupC	p.Cys665LeufsTer13	16		null allele	het	9	F	B+	12	0	5	3.6	0
			c.4146G>T	c.4146G>T (p.=)	28		A1	het								
			c.6197A>G	p.Asn2066Ser	36		D4	het								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	het								
0509	05010	3 (carrier)	c.3735G>A	c.3735G>A (p.=)	28		D3	het	12	F	B+	4	37	40.6	69	41
			c.3797C>T	p.Pro1266Leu	28		D3	hom								
			c.6187C>T	**p.Pro2063Ser**	36		D4	het								
0510	05011*	3	c.5837G>T	p.Cys1946Phe	34		D4	het	9	F	A+	13	0	5	3.3	0
0514	05015	3 (carrier)	c.1533+15G>A			13	intronic	het	43	F	O+	4	59	60	109	54
0515	05016	3	c.100C>T	**p.Arg34Ter**	3		null allele	hom	47	F	B-	13	0	5	12	0
	05017	3	c.100C>T	**p.Arg34Ter**	3		null allele	hom	47	F	B-	18	0	5	4.1	0
0803	08003	3 (carrier)	c.6187C>T	**p.Pro2063Ser**	36		D4	het	23	F	O+	6	76	78	108	80
	08004	3 (carrier)	c.6187C>T	**p.Pro2063Ser**	36		D4	het	16	F	A+	6	33	39	75	40
1001	10001	3	c.7664_7665insAG	**p.Cys2557SerfsTer8**	45		null allele	hom	53	M	O+	2	5.1	5	13	4.3
1004	10005	3	c.1534_1536del	p.Leu512del	14		D2	het	26	F	O-	4	0	5	3.5	0
			c.4146G>T	c.4146G>T (p.=)	28		A1	het								
			c.6197A>G	p.Asn2066Ser	36		D4	het								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	het								
1005	10006*	3	c.7408C>T	p.Gln2470Ter	43		null allele	hom	14	M	B+	8	68	29	59	41
1007	10008	3	c.1093C>T	**p.Arg365Ter**	9		null allele	hom	29	M	O+	2	3.2	5	10	2.5
1009	10011*	3	c.5455+1G>A			31	intronic	het	33	F	O+	1	0	5	4.4	0
1014	10016	3	c.7118C>T	p.Pro2373Leu	42		C2	het	19	M	A+	4	0	5	3	0
			c.7408C>T	p.Gln2470Ter	43		null allele	het								
			c.8347C>T	p.Gln2783Ter	52		null allele	het								
1217	12024	3 (carrier)	c.6187C>T	**p.Pro2063Ser**	36		D4	het	31	M	O+	0	56	40	59	41
1306	13008	3	c.4146G>T	c.4146G>T (p.=)	28		A1	hom	39	F	A+	18	3	4	2.7	0.7
			c.6197A>G	p.Asn2066Ser	36		D4	hom								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	hom								
2001	20001	3	c.4146G>T	c.4146G>T (p.=)	28		A1	hom	9	F	O-	6	0	5	1.6	0
			c.6197A>G	p.Asn2066Ser	36		D4	hom								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	hom								
	20003	3 (carrier)	c.4146G>T	c.4146G>T (p.=)	28		A1	het	46	M	A+	0	45	38	63	34
			c.6197A>G	p.Asn2066Ser	36		D4	het								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	het								
2101	21001	3	c.4162C>T	p.Gln1388Ter	28		null allele	hom	45	M	ND	14	0	5	2.5	0
2102	21002	3	c.3426T>C	c.3426T>C (p.=)	26		D3	hom	41	F	B+	12	0	5	6.8	0
			c.3485_3486delinsTG	p.Pro1162Leu	26		D3	hom								
2104	21007	3	c.5311G>A	p.Gly1771Arg	30		A3	het	51	F	A+	16	2.3	5	10	0
			c.818G>C	p.Arg273Pro	7		D1	het								
2113	21015	3	c.4162C>T	p.Gln1388Ter	28		null allele	het	39	M	A+	16	0	5	6.4	0
			c.7583G>A	p.Cys2528Tyr	45		C4	het								

Family	Patient	Phenotype	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:CB
2505	25005	3	c.(1945+1_1946-1)_(7437+1_7438-1)del		16-43		D2-C3	hom	32	M	O+	11	0	4	2.4	0
2716	27034	3 (carrier)	c.2518G>T	p.Glu840Ter	19		null allele	het	16	F	A+	3	42	32	78	54
2720	27038	3 (carrier)	c.6187C>T	**p.Pro2063Ser**	36		D4	het	51	M	A+	-1	68	52	67	59
2721	27041	3 (carrier)	c.2944G>C	p.Val982Leu	22		D3	het	43	F	A+	4	42	25	66	48
			c.4146G>T	c.4146G>T (p.=)	28		A1	het								
			c.6197A>G	p.Asn2066Ser	36		D4	het								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	het								
3001	30001	3	c.4975C>T	**p.Arg1659Ter**	28		null allele	het	60	F	O+	26	0	5	2.2	0
			c.8347C>T	p.Gln2783Ter	52		null allele	het								
3210	32011	3 (carrier)	c.3931C>T	**p.Gln1311Ter**	28		null allele	het	31	F	A+	3	51	23.5	88	21
3401	34001	3 (carrier)	c.1209C>G		11		null allele	het	42	M	O-	3	41	35	48	28
			c.2821-88A>G			21	intronic	het								
			c.425G>A	p.Gly142Asp	5		D1	het								
			c.7771-82G>A			46	intronic	het								
3605	36007	3	c.2771G>A	p.Arg924Gln	21		D3	het	52	M	A-	10	4.1	5	5.7	0
			c.2900G>T	p.Gly967Val	22		D3	het								
			c.6799-27C>T			38	intronic	het								
3806	38008	3	c.2025G>A	c.2025G>A (p.=)	16		D2	het	12	F	ND	11	4.3	4	13	3.7
			c.3797C>T	p.Pro1266Leu	28		D3	het								
			c.3835G>A	p.Val1279Ile	28		A1	het								
			c.5198T>C	p.Leu1733Pro	30		A3	het								
3905	39009	3	c.6598+1G>A			37	intronic	het	37	M	B-	19	0	4	2.3	0
			c.8347C>T	p.Gln2783Ter	52		null allele	het								
	39010	3 (carrier)	c.6598+1G>A			37	intronic	het	76	F	B-	-3	104	92	128	57
	39014	3	c.6598+1G>A			37	intronic	het	37	M	B-	20	0	4	2.5	0
			c.8347C>T	p.Gln2783Ter	52		null allele	het								
4401	44001	3	c.4146G>T	c.4146G>T (p.=)	28		A1	hom	23	M	ND	21	0	4	1.3	0
			c.6197A>G	p.Asn2066Ser	36		D4	hom								
			c.375_376delinsC	**p.Tyr126ThrfsTer49**	5		null allele	hom								

In bold, mutations previously described in type 3. ND indicates not determined.* Patients with a discrepant phenotype-genotype correlation.

Table S5. Genotype-phenotype correlation in type 2A patients

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
0101	01001	c.3445T>C	**p.Cys1149Arg**	26		D3	het	44	F	B+	7	↓HMWM	10	6.8	18	0.68	5.8
	01045	c.3445T>C	**p.Cys1149Arg**	26		D3	het	49	F	B+	12	↓HMWM	13	7.7	20	0.59	8.7
		c.7070A>T	p.Asn2357Ile	41		C2	het										
0105	01008	c.4121G>A	**p.Arg1374His**	28		A1	het	68	F	O+	8	↓HMWM	20	4	45	0.20	13
		c.-1875G>A				upstream	het										
0108	01012	c.1946-17_1946-16insTTT			15	intronic	het	45	F	B+	16	↓HMWM	16	4	36	0.25	1.6
		c.3815G>C	**p.Cys1272Ser**	28		A1	het										
0110	01018	c.4196G>A	p.Arg1399His	28		A1	het	31	F	A-	12	↓HMWM	8.3	4	18	0.48	3.4
		c.4517C>T	**p.Ser1506Leu**	28		A2	het										
0123	01043	c.3445T>C	**p.Cys1149Arg**	26		D3	het	25	F	B+	7	↓HMWM	9.5	4	11.6	0.42	6.3
	01046	c.3445T>C	**p.Cys1149Arg**	26		D3	het	65	M	A+	9	↓HMWM	8	5	18	0.63	5
	01047	c.3445T>C	**p.Cys1149Arg**	26		D3	het	40	M	A+	12	↓HMWM	8.6	5.2	14	0.60	4.7
	01053	c.3445T>C	**p.Cys1149Arg**	26		D3	het	36	M	A+	10	↓HMWM	15	8.1	20	0.54	6.1
0125	01048	c.3445T>C	**p.Cys1149Arg**	26		D3	het	25	F	B+	18	↓HMWM	11	7.6	15	0.69	5.1
	01061	c.3445T>C	**p.Cys1149Arg**	26		D3	het	53	M	B+	2	↓HMWM	17	7.9	27	0.46	8.2
0127	01069	c.4517C>T	**p.Ser1506Leu**	28		A2	het	17	F	O+	13	↓HMWM	19	4	30	0.21	6.3
		c.7344C>T	c.7344C>T (p.=)	43		C3	het										
	01070	c.4517C>T	**p.Ser1506Leu**	28		A2	het	38	F	O+	21	↓HMWM	38	7.2	59	0.19	12
0130	01055	c.3445T>C	**p.Cys1149Arg**	26		D3	het	18	F	O+	6	↓HMWM	12	7.9	20	0.66	5.2
	01056	c.3445T>C	**p.Cys1149Arg**	26		D3	het	38	F	A+	8	↓HMWM	13	7.3	23	0.56	6.1
	01057	c.3445T>C	**p.Cys1149Arg**	26		D3	het	71	F	O+	10	↓HMWM	9.6	4	13	0.42	4.4
0135	01072	c.4789C>T	**p.Arg1597Trp**	28		A2	het	48	M	O	7	↓HMWM	37	22	51	0.59	15
0203	02010	c.4789C>T	**p.Arg1597Trp**	28		A2	het	54	F	O+	21	↓↓HMWM	25	6.5	91	0.26	9.4
	02011	c.4789C>T	**p.Arg1597Trp**	28		A2	het	31	F	O+	13	↓↓HMWM	27	8	122	0.30	8.9
	02012	c.4789C>T	**p.Arg1597Trp**	28		A2	het	27	F	ND	8	↓↓HMWM	33	10.2	45	0.31	12
	02013	c.4789C>T	**p.Arg1597Trp**	28		A2	het	18	M	ND	11	↓↓HMWM	16	8.4	49	0.53	8.1
0205	02021	c.4883T>C	**p.Ile1628Thr**	28		A2	het	44	F	A-	17	↓↓HMWM	38	6	61	0.16	6.6
	02022	c.4883T>C	**p.Ile1628Thr**	28		A2	het	14	F	A-	8	↓↓HMWM	43	7.3	64	0.17	7.4
		c.5191T>A	p.Ser1731Thr	30		A3	het										
0209	02028	c.4825G>A	**p.Gly1609Arg**	28		A2	het	47	F	A-	18	↓↓HMWM	35	7.8	67	0.22	8.6
		c.6187C>T	p.Pro2063Ser	36		D4	het										
	02029	c.4825G>A	**p.Gly1609Arg**	28		A2	het	14	F	A+	7	↓↓HMWM	31	7.2	46	0.23	13
		c.6187C>T	p.Pro2063Ser	36		D4	het										

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
0227	02071*	c.2926C>T	p.Arg976Cys	22		D3	het	34	F	O+	3	NORMAL	27	23.6	67	0.87	36
		c.6187C>T	p.Pro2063Ser	36		D4	het										
	02072*	c.2926C>T	p.Arg976Cys	22		D3	het	62	F	O+	5	NORMAL	45	46	114	1.02	57
		c.6187C>T	p.Pro2063Ser	36		D4	het										
0231	02081	c.4885G>A	**p.Gly1629Arg**	28		A2	het	43	M	O-	9	↓↓HMWM	35	6.4	30	0.18	15
0304	03007	c.3426T>C	c.3426T>C (p.=)	26		D3	het	50	M	A+	8	↓HMWM	18	6.4	30	0.36	15
		c.4121G>A	**p.Arg1374His**	28		A1	het										
		c.8156-42C>T			50	intronic	het										
		c.3485_3486delinsTG	p.Pro1162Leu	26		D3	het										
	03008	c.1156+42C>T			10	intronic	het	23	F	ND	3	↓HMWM	14	7	26	0.50	14
		c.4121G>A	**p.Arg1374His**	28		A1	het										
03152	03012*	c.1157_5620del	p.Gly386_Ser1873del	11-32		D2-D3	het	62	F	O+	10	↓HMWM	42	20	70	0.48	21
03231	03048	c.3426T>C	c.3426T>C (p.=)	26		D3	het	15	F	O+	0	↓HMWM	117	6.5	90	0.06	30.8
		c.8318G>C	**p.Cys2773Ser**	52		CK	het										
		c.3485_3486delinsTG	p.Pro1162Leu	26		D3	het										
0329	03026	c.4883T>C	**p.Ile1628Thr**	28		A2	het	73	M	A+	8	↓HMWM	56	9.1	29	0.16	11
	03027	c.4883T>C	**p.Ile1628Thr**	28		A2	het	42	F	O+	5	↓HMWM	41	8.7	29	0.21	7
	03028	c.4883T>C	**p.Ile1628Thr**	28		A2	het	9	M	O+	10	↓HMWM	38	7.4	22	0.19	7.7
	03029	c.4883T>C	**p.Ile1628Thr**	28		A2	hom	44	F	O+	6	↓HMWM	47	6	32	0.13	8.9
0336	03009	c.4678_4680dup	p.Asp1560dup	28		A2	het	45	F	O+	5	↓HMWM	19	5	35	0.26	13
	03010	c.4678_4680dup	p.Asp1560dup	28		A2	het	25	F	ND	5	↓HMWM	49	5.3	44	0.11	18
0347	03011	c.1625C>G	p.Ala542Gly	14		D2	het	80	M	A+	4	↓HMWM	41	12.1	89	0.30	15
		c.1721C>G	p.Pro574Arg	14		D2	het										
		c.533-2A>G			5	intronic	het										
0351	03040	c.4883T>C	**p.Ile1628Thr**	28		A2	het	35	F	A+	7	↓HMWM	43	6.3	39	0.15	8.2
		c.7344C>T	c.7344C>T (p.=)	43		C3	het										
0505	05005	c.4889T>G	p.Val1630Gly	28		A2	het	47	M	AB+	8	↓HMWM	19	5	36	0.26	8.5
	05006	c.4889T>G	p.Val1630Gly	28		A2	het	9	F	B-	4	↓HMWM	16	6.2	43	0.39	6.8
0615	06021	c.4121G>A	**p.Arg1374His**	28		A1	het	36	F	O+	3	↓HMWM	13	5.7	31	0.44	13
		c.6187C>T	p.Pro2063Ser	36		D4	het										
	06022	c.4121G>A	**p.Arg1374His**	28		A1	het	11	M	O-	0	↓HMWM	16	6.1	30	0.38	16
0702	07002	c.4840G>A	p.Asp1614Asn	28		A2	het	39	F	O+	15	↓HMWM	18	6.4	24	0.36	8.9
	07005	c.4840G>A	p.Asp1614Asn	28		A2	het	13	F	O+	11	↓HMWM	19	8.3	27	0.44	13
	07006	c.4840G>A	p.Asp1614Asn	28		A2	het	13	F	O+	11	↓HMWM	17	4	26	0.24	12
0804	08005	c.4499C>A	**p.Ala1500Glu**	28		A2	het	8	F	A+	9	↓↓HMWM	15	5	34	0.33	9.4
		c.4620G>C	c.4620G>C (p.=)	28		A2	het										

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
1006	10007	c.4883T>C	**p.Ile1628Thr**	28		A2	het	49	F	A+	1	↓↓HMWM	56	5	59	0.09	11
	10009	c.4883T>C	**p.Ile1628Thr**	28		A2	het	43	M	A+	3	↓↓HMWM	48	5	47	0.10	9.9
1303	13004	c.4883T>C	**p.Ile1628Thr**	28		A2	het	42	M	A+	17	↓↓HMWM	26	5	31	0.19	7
1310	13013	c.4517C>T	**p.Ser1506Leu**	28		A2	het	37	M	A+	20	↓↓HMWM	18	5	29	0.28	7.2
1312	13015	c.4622A>T	p.Gln1541Leu	28		A2	het	48	F	B+	22	↓↓HMWM	30	9.3	43	0.31	10.5
		c.6182del	p.Phe2061SerfsTer38	36		D4	het										
1404	14004	c.4825G>A	**p.Gly1609Arg**	28		A2	het	63	F	B+	9	↓↓HMWM	78	15	62	0.19	18
1405	14005	c.4789C>T	**p.Arg1597Trp**	28		A2	het	10	F	O+	11	↓↓HMWM	24	5	35	0.21	6.6
2002	20002	c.4517C>T	**p.Ser1506Leu**	28		A2	het	22	F	O-	13	↓HMWM	12	6.1	19	0.51	6
2116	21018	c.250C>T	p.Leu84Phe	4		D1	het	41	F	B+	14	↓HMWM	16	6.2	44	0.39	11
		c.4739T>C	**p.Leu1580Pro**	28		A2	het										
		c.6187C>T	p.Pro2063Ser	36		D4	het										
2202	22002	c.4960T>C	**p.Phe1654Leu**	28		A2	het	19	F	A+	9	↓HMWM	35	7.3	50	0.21	18
		c.5368C>T	p.Pro1790Ser	31		A3	het										
2203	22003	c.4790G>A	**p.Arg1597Gln**	28		A2	het	72	F	A+	3	↓HMWM	34	7.5	66	0.22	16
		c.6976+111G>A			40	intronic	het										
	22008	c.4790G>A	**p.Arg1597Gln**	28		A2	het	70	M	ND	11	↓HMWM	39	6.8	40	0.17	13
		c.6976+111G>A			40	intronic	het										
	22009	c.4790G>A	**p.Arg1597Gln**	28		A2	het	41	F	ND	14	↓HMWM	53	8.2	48	0.15	17
	22010	c.4790G>A	**p.Arg1597Gln**	28		A2	het	38	M	A+	10	↓HMWM	66	10.6	51	0.16	22
2205	22005	c.4241T>G	p.Val1414Gly	28		A1	het	31	F	O+	10	↓HMWM	12	6.9	47	0.58	11
2301		c.2821-123A>C			21	intronic	het										
	23001	c.4960T>C	**p.Phe1654Leu**	28		A2	het	19	F	A+	11	↓HMWM	48	6.5	43	0.14	18
		c.6187C>T	p.Pro2063Ser	36		D4	het										
	23002	c.4960T>C	**p.Phe1654Leu**	28		A2	het	42	F	A+	14	↓HMWM	66	7.1	44	0.11	20
	23014	c.1446C>G	p.Ile482Met	13		D2	het	21	M	A+	4	↓HMWM	82	7.2	45	0.09	24
		c.4257T>G	p.His1419Gln	28		A1	het										
		c.4960T>C	**p.Phe1654Leu**	28		A2	het										
2702	27003	c.2771G>A	p.Arg924Gln	21		D3	het	75	M	ND	9	↓↓↓HMWM	40	9.9	43	0.25	10
		c.4943C>G	**p.Pro1648Arg**	28		A2	het										
2705	27008	c.4735G>A	**p.Gly1579Arg**	28		A2	het	56	F	A+	6	↓↓HMWM	43	9	45	0.21	10
3002	30003	c.4517C>T	**p.Ser1506Leu**	28		A2	het	49	M	A+	10	↓↓HMWM	18	5.9	16	0.33	7.1
		c.7393G>A	p.Val2465Met	43		C3	het										

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
3003	30004	c.2821-78A>C			21	intronic	het	30	F	B+	6	↓HMWM	16	5.4	26	0.34	13
		c.3920T>C	**p.Leu1307Pro**	28		A1	het										
	30007	c.3920T>C	**p.Leu1307Pro**	28		A1	het	61	F	B+	22	↓HMWM	13	5.9	24	0.45	9.9
		c.-1650G>C				upstream	het										
	30010	c.3920T>C	**p.Leu1307Pro**	28		A1	het	66	M	A+	13	↓HMWM	11	5.6	21	0.51	9.2
3004	30005	c.4121G>A	**p.Arg1374His**	28		A1	het	67	F	O+	20	NORMAL	22	5.8	26	0.26	15
		c.5277C>T	c.5277C>T (p.=)	30		A3	het										
	30011	c.2685+147G>A			20	intronic	het	38	M	O-	9	NORMAL	18	5.8	26	0.32	15
		c.4121G>A	**p.Arg1374His**	28		A1	het										
	30014	c.4121G>A	**p.Arg1374His**	28		A1	het	49	M	O+	16	↓HMWM	15	4	14	0.27	6.7
	30015	c.4121G>A	**p.Arg1374His**	28		A1	het	18	M	A+	2	NORMAL	31	5	29	0.16	14
	30016	c.2821-148G>A			21	intronic	het	14	M	O-	9	NORMAL	22	5	26	0.23	12
		c.4121G>A	**p.Arg1374His**	28		A1	het										
	30019	c.4121G>A	**p.Arg1374His**	28		A1	het	75	F	O+	15	NORMAL	26	5	21	0.19	12
	30020	c.4121G>A	**p.Arg1374His**	28		A1	het	71	F	O+	11	NORMAL	62	5.6	45	0.09	27
		c.5277C>T	c.5277C>T (p.=)	30		A3	het										
	30021	c.4121G>A	**p.Arg1374His**	28		A1	het	60	F	A+	14	NORMAL	35	5.8	33	0.17	17
		c.5277C>T	c.5277C>T (p.=)	30		A3	het										
	30022	c.4121G>A	**p.Arg1374His**	28		A1	het	62	F	O+	16	NORMAL	30	5.6	25	0.19	14
	30024	c.4121G>A	**p.Arg1374His**	28		A1	het	18	M	A+	7	NORMAL	25	5.4	19	0.22	14
	30025	c.4121G>A	**p.Arg1374His**	28		A1	het	47	M	A+	12	NORMAL	26	4	25	0.15	11
	30026	c.4121G>A	**p.Arg1374His**	28		A1	het	37	F	A+	10	NORMAL	33	4	27	0.12	18
	30029	c.4121G>A	**p.Arg1374His**	28		A1	het	48	F	A+	10	NORMAL	19	5.7	19	0.30	13
	30031	c.4121G>A	**p.Arg1374His**	28		A1	het	36	F	A+	8	NORMAL	22	4	21	0.18	11
	30032	c.4121G>A	**p.Arg1374His**	28		A1	het	69	F	A+	6	NORMAL	57	9	41	0.16	29
	30034	c.4121G>A	**p.Arg1374His**	28		A1	het	8	F	A+	12	NORMAL	21	5	21	0.24	9.8
	30036	c.4121G>A	**p.Arg1374His**	28		A1	het	24	M	O+	2	NORMAL	23	5	30	0.22	11
	30038	c.4121G>A	**p.Arg1374His**	28		A1	het	24	F	A+	6	↓HMWM	29	5	24	0.17	12
	30039	c.4121G>A	**p.Arg1374His**	28		A1	het	33	M	A+	3	↓HMWM	21	5	20	0.24	7.5
	30040	c.4121G>A	**p.Arg1374His**	28		A1	het	33	F	A+	10	↓HMWM	22	6.9	33	0.31	11
3009	30017	c.2821-123A>C			21	intronic	het	17	M	O-	7	↓↓HMWM	17	5	12	0.29	4.3
		c.4789C>T	**p.Arg1597Trp**	28		A2	het										
		c.5001G>A	c.5001G>A (p.=)	28		A2	het										
3701	37001	c.4789C>G	**p.Arg1597Gly**	28		A2	het	36	F	A+	15	↓↓HMWM	41	5	40	0.12	15
	37007	c.4789C>G	**p.Arg1597Gly**	28		A2	het	3	F	AB+	3	↓↓HMWM	48	8.8	33	0.18	21
3708	37009	c.4825G>A	**p.Gly1609Arg**	28		A2	het	41	F	O-	6	↓HMWM	60	6.6	29	0.11	18
3802	38002	c.4883T>C	**p.Ile1628Thr**	28		A2	het	13	M	O+	23	↓HMWM	57	4	50	0.07	15

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
3902	39001	c.3433C>T	p.Arg1145Cys	26		D3	het	70	M	O+	10	↓HMWM	18	5	35	0.28	17
		c.4121G>A	**p.Arg1374His**	28		A1	het										
		c.-1875G>A				upstream	het										
	39004	c.4121G>A	**p.Arg1374His**	28		A1	het	33	M	A+	11	ND	13	6.7	29	0.52	15
		c.-1875G>A				upstream	het										
3903	39005	c.4121G>A	**p.Arg1374His**	28		A1	het	35	M	A-	8	↓HMWM	13	7.1	26	0.55	14
		c.-2076A>G				upstream	het										
	39006	c.4121G>A	**p.Arg1374His**	28		A1	het	59	F	A+	15	↓HMWM	12	4.7	23	0.39	14
3906	39011	c.4517C>T	**p.Ser1506Leu**	28		A2	het	62	F	A+	11	↓↓HMWM	17	7.7	31	0.45	21
	39012	c.4517C>T	**p.Ser1506Leu**	28		A2	het	41	M	O+	15	↓↓HMWM	8.2	5	13	0.61	6.9
	39013	c.4517C>T	**p.Ser1506Leu**	28		A2	het	31	M	B+	12	↓↓HMWM	21	6.8	30	0.32	11
3907	39016	c.4517C>T	**p.Ser1506Leu**	28		A2	het	17	F	A-	11	↓↓HMWM	37	7.7	36	0.21	14
		c.5277C>T	c.5277C>T (p.=)	30		A3	het										
4201	42001	c.4121G>A	**p.Arg1374His**	28		A1	het	70	M	O+	8	↓HMWM	27	4	24	0.15	14
4210	42010*	c.4580G>A	**p.Arg1527Gln**	28		A2	het	9	M	O+	2	↓HMWM	69	58.5	52	0.85	45
4403	44004	c.4892G>A	p.Gly1631Asp	28		A2	het	19	F	A+	14	↓HMWM	51	5	54	0.10	12
4502	45002	c.4789C>T	**p.Arg1597Trp**	28		A2	het	16	F	ND	6	↓HMWM	22	5	27	0.23	8
	45003	c.4789C>T	**p.Arg1597Trp**	28		A2	het	48	M	O+	9	↓HMWM	26	5	30	0.19	9.3
4503	45004	c.3944G>T	**p.Arg1315Leu**	28		A1	het	49	M	ND	11	↓HMWM	21	5	31	0.24	14

In bold, mutations previously described in type 2A. ND indicates not determined; and HMWM, high molecular weight multimers. *Patients with a discrepant phenotype-genotype correlation.

Table S6. Genotype-phenotype correlation in type 2A/2M patients

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
0102	01002	c.4120C>T	**p.Arg1374Cys**	28		A1	het	51	F	O+	11	SMEAR	16	4.4	34	0.28	9.2
		c.6187C>T	p.Pro2063Ser	36		D4	het										
	01003	c.4120C>T	**p.Arg1374Cys**	28		A1	het	58	F	A	15	SMEAR	20	6.7	45	0.34	13
	01014	c.4120C>T	**p.Arg1374Cys**	28		A1	het	55	F	O	19	SMEAR	15	4.5	29	0.30	11
	01020	c.4120C>T	**p.Arg1374Cys**	28		A1	het	66	F	O+	12	SMEAR	32	11	47	0.34	19
		c.6433C>T	p.Pro2145Ser	37		D4	het										
	01022	c.4120C>T	**p.Arg1374Cys**	28		A1	het	80	M	A	11	SMEAR	54	18	59	0.33	30
	01035	c.4120C>T	**p.Arg1374Cys**	28		A1	het	38	F	A-	18	SMEAR	14	6.7	25	0.48	15
	01039	c.2771G>A	p.Arg924Gln	21		D3	het	18	M	A+	8	SMEAR	23	12	24	0.52	15
		c.4120C>T	**p.Arg1374Cys**	28		A1	het										
0104	01007	c.4120C>T	**p.Arg1374Cys**	28		A1	het	66	M	AB+	6	SMEAR	25	7.8	38	0.31	15
	01009	c.3835G>A	p.Val1279Ile	28		A1	het	88	F	A+	9	SMEAR	17	4	52	0.24	12
		c.4120C>T	**p.Arg1374Cys**	28		A1	het										
	01015	c.4120C>T	**p.Arg1374Cys**	28		A1	het	39	F	AB	12	SMEAR	19	7.8	36	0.41	11
	01023	c.4120C>T	**p.Arg1374Cys**	28		A1	het	38	M	AB+	7	SMEAR	30	13	27	0.43	18
	01026	c.3835G>A	p.Val1279Ile	28		A1	het	90	M	A+	16	SMEAR	18	4	37	0.22	18
		c.4120C>T	**p.Arg1374Cys**	28		A1	het										
	01027	c.4120C>T	**p.Arg1374Cys**	28		A1	het	11	F	B+	8	SMEAR	11	6.4	28	0.58	10
		c.4255C>A	p.His1419Asn	28		A1	het										
	01029	c.4120C>T	**p.Arg1374Cys**	28		A1	het	60	F	AB+	10	SMEAR	18	10	34	0.56	19
		c.6798+32G>A			38	intronic	het										
	01031	c.4120C>T	**p.Arg1374Cys**	28		A1	het	16	M	A+	15	SMEAR	12.3	5.5	21	0.45	10
		c.4255C>A	p.His1419Asn	28		A1	het										
	01033	c.4120C>T	**p.Arg1374Cys**	28		A1	het	19	F	B+	10	SMEAR	27	10	32	0.37	16
0112	01024	c.4120C>T	**p.Arg1374Cys**	28		A1	het	33	F	O	10	SMEAR	12	4	24	0.33	11
		c.7730-4C>G			45	intronic	het										
0116	01032	c.4120C>T	**p.Arg1374Cys**	28		A1	het	10	M	A+	6	SMEAR	15.7	4.5	19	0.29	11
		c.7393G>A	p.Val2465Met	43		C3	het										
	01042	c.4120C>T	**p.Arg1374Cys**	28		A1	het	79	F	O+	0	SMEAR	27	4.9	26	0.18	17
	01071	c.4120C>T	**p.Arg1374Cys**	28		A1	het	50	M	O+	12	SMEAR	13	6.6	17	0.51	6.7
		c.4196G>A	p.Arg1399His	28		A1	het										
0120	01038	c.3943C>T	**p.Arg1315Cys**	28		A1	het	59	F	O+	7	SMEAR	20	5.3	20	0.27	13
		c.5842+31C>T			34	intronic	het										
		c.-2627C>T				upstream	het										
0208	02026	c.3943C>T	**p.Arg1315Cys**	28		A1	het	55	M	O-	10	SMEAR	21	7.4	44	0.35	8.8
		c.-1896C>T				upstream	het										
	02027	c.3943C>T	**p.Arg1315Cys**	28		A1	het	31	M	O+	16	SMEAR	11	7.1	28	0.65	6.4
		c.-1896C>T				upstream	het										

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
0212	02034	c.2446C>T	p.Arg816Trp	19		D'	het	63	F	O+	16	SMEAR	26	8.4	23	0.32	14
		c.3943C>T	**p.Arg1315Cys**	28		A1	het										
	02035	c.3943C>T	**p.Arg1315Cys**	28		A1	het	35	M	O+	5	SMEAR	12	6.5	31	0.54	9.2
0376	03017	c.4120C>T	**p.Arg1374Cys**	28		A1	het	65	M	O+	9	SMEAR	44	11.8	39	0.27	23
1307	13009	c.3943C>T	**p.Arg1315Cys**	28		A1	het	40	F	O+	19	SMEAR	9.5	5	20	0.53	5.8
	13016	c.3943C>T	**p.Arg1315Cys**	28		A1	het	11	M	ND	8	SMEAR	11	5	21	0.45	6.7
3007	30012*	c.2289G>C	p.Arg763Ser	18		D2	het	33	F	O-	15	SMEAR	32	28	18	0.88	19
	30030*	c.2289G>C	p.Arg763Ser	18		D2	het	53	M	O-	6	SMEAR	40	25.4	22	0.64	29
3008	30013*	c.2289G>C	p.Arg763Ser	18		D2	het	79	M	A+	4	SMEAR	78	31.8	53	0.41	34
3703	37003	c.7471T>C	p.Cys2491Arg	44		C3	het	31	F	O+	0	SMEAR	30	25	44	0.83	20
	37011	c.7471T>C	p.Cys2491Arg	44		C3	het	37	F	O+	7	SMEAR	34	17.7	43	0.52	24
3904	39008	c.4120C>T	**p.Arg1374Cys**	28		A1	het	73	M	A+	6	SMEAR	15	7.7	33	0.51	17
		c.7771-40G>A			46	intronic	het										

In bold, mutations included in EAHAD-VWD-LOVD. ND indicates not determined. *Patients that present an altered VWF:FVIII:B 30012=0.28; 30003=0.37 and 30013=0.42 leading to a slightly 2N phenotype due to the presence of the new mutation p.Arg763Ser.

Table S7. Genotype-phenotype correlation in type 2B patients

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
0134	01065	c.3922C>T	**p.Arg1308Cys**	28		A1	het	40	F	A+	13	↓HMWM	63	19	50	0.30	23
0207	02024	c.3922C>T	**p.Arg1308Cys**	28		A1	het	53	F	O-	7	↓HMWM	21	9	37	0.43	6.3
	02025	c.2771G>A	p.Arg924Gln	21		D3	het	30	M	O+	7	↓HMWM	25	8.8	33	0.35	5.7
		c.3922C>T	**p.Arg1308Cys**	28		A1	het										
	02078	c.3922C>T	**p.Arg1308Cys**	28		A1	het	76	F	AB-	16	↓HMWM	34	24	55	0.71	21
0210	02030	c.4010C>T	**p.Pro1337Leu**	28		A1	het	50	M	A+	15	↓HMWM	75	35.4	134	0.47	50
		c.6237A>G	c.6237A>G (p.=)	36		D4	het										
0213	02036	c.3797C>T	**p.Pro1266Leu**	28		D3	het	34	F	O+	10	NORMAL	31	33	45	1.06	28
	02037	c.3797C>T	**p.Pro1266Leu**	28		D3	het	58	F	O-		NORMAL	67	65	114	0.97	67
0214	02038	c.3916C>T	**p.Arg1306Trp**	28		A1	het	45	F	A+	12	↓HMWM	31	21	75	0.68	18
		c.8366C>G	p.Thr2789Ser	52		CK	het										
0516	05018	c.3916C>T	**p.Arg1306Trp**	28		A1	het	82	M	A+	13	↓HMWM	119	72.8	63	0.61	71
0611	06015	c.3916C>T	**p.Arg1306Trp**	28		A1	het	67	M	O	11	↓HMWM	42	23	60	0.55	27
	06016	c.3379+27A>T			25	intronic	het	40	F	O-	8	↓HMWM	49	31	58	0.63	32
		c.3916C>T	**p.Arg1306Trp**	28		A1	het										
		c.7548+22G>A			44	intronic	het										
0614	06020	c.3922C>T	**p.Arg1308Cys**	28		A1	het	28	F	A+	20	↓HMWM	66	18	55	0.27	22
0806	08006	c.3916C>T	**p.Arg1306Trp**	28		A1	het	7	F	A+	11	↓↓HMWM	27	5	30	0.19	9.1
1012	10014	c.3946G>A	**p.Val1316Met**	28		A1	het	53	F	O+	4	↓HMWM	40	18	50	0.45	14
1208	12015	c.3917G>A	**p.Arg1306Gln**	28		A1	het	61	F	O+	13	↓HMWM	56	45	48	0.80	42
1302	13002	c.3922C>T	**p.Arg1308Cys**	28		A1	het	23	M	A+	11	↓↓HMWM	42	13.2	38	0.31	15
	13003	c.3922C>T	**p.Arg1308Cys**	28		A1	het	54	M	A+	8	↓↓HMWM	39	12.2	49	0.31	17
	13005	c.3922C>T	**p.Arg1308Cys**	28		A1	het	61	M	O+	8	↓↓HMWM	26	8.6	42	0.33	14
1305	13007	c.3922C>T	**p.Arg1308Cys**	28		A1	het	41	M	A+	24	↓↓HMWM	63	14	33	0.22	15
		c.-2692C>T				upstream	het										
1313	13017	c.3922C>T	**p.Arg1308Cys**	28		A1	het	50	M	O+	3	↓↓HMWM	21	14	30	0.67	0.61
1401	14001	c.3916C>T	**p.Arg1306Trp**	28		A1	het	28	M	O+	10	↓↓HMWM	9	6	15	0.67	7.6
1905	19007	c.1432+166C>A			12	intronic	het	38	F	O+	15	↓HMWM	32	18.7	32	0.58	17
		c.3426T>C	c.3426T>C (p.=)	26		D3	het										
		c.3946G>A	**p.Val1316Met**	28		A1	het										
		c.5278G>A	p.Val1760Ile	30		A3	het										
		c.3485_3486delinsTG	p.Pro1162Leu	26		D3	het										
2119	21022	c.3916C>T	**p.Arg1306Trp**	28		A1	het	42	F	A+	4	↓HMWM	44	20.5	37	0.47	19
2708	27013	c.3916C>T	**p.Arg1306Trp**	28		A1	het	59	F	O-	11	↓↓HMWM	22	10.3	45	0.47	12
	27014	c.3916C>T	**p.Arg1306Trp**	28		A1	het	32	M	O-	8	↓↓HMWM	25	10.8	38	0.43	14
	27015	c.3916C>T	**p.Arg1306Trp**	28		A1	het	67	M	O-	3	↓↓HMWM	33	18.5	63	0.56	25
	27016	c.3916C>T	**p.Arg1306Trp**	28		A1	het	62	F	O-	13	↓↓HMWM	29	14.1	40	0.49	19
		c.5277C>T	c.5277C>T (p.=)	30		A3	het										
3203	32003	c.3946G>A	**p.Val1316Met**	28		A1	het	28	M	B+	13	↓↓HMWM	36	12.2	32	0.34	15
		c.4751A>G	p.Tyr1584Cys	28		A2	het										
3208	32008	c.3917G>A	**p.Arg1306Gln**	28		A1	het	34	F	AB+	9	↓HMWM	116	46	50	0.40	44

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB
3504	35009	c.3922C>T	**p.Arg1308Cys**	28		A1	het	40	F	O	10	↓↓HMWM	23	18	24	0.78	16
	35010	c.3922C>T	**p.Arg1308Cys**	28		A1	het	10	M	O-	5	↓↓HMWM	25	14	41	0.56	12
3602*	36002	c.3946G>A	**p.Val1316Met**	28		A1	het	58	F	O+	14	NORMAL*	57	21	70	0.37	30
	36003	c.3946G>A	**p.Val1316Met**	28		A1	het	37	M	O+	12	↓HMWM*	58	23	97	0.40	40
4101	41001	c.3797C>A	**p.Pro1266Gln**	28		D3	het	17	M	O+	3	NORMAL	41	36	48	0.88	30
	41002	c.3797C>A	**p.Pro1266Gln**	28		D3	het	20	M	O+	0	NORMAL	39	31	41	0.79	30

In bold, mutations previously described in type 2B. ND indicates not determined.; and HMWM, high molecular weight multimers. * The family 3602 had the mutation p.Val1316Met related with Montreal platelet syndrome, detected in both related patients, but results in different multimeric patterns, as previoulsy described by Rendal *et al.*

Table S8. Genotype-phenotype correlation in type 2M patients

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB	VWF:CB(VI)
0109	01013	c.220+100G>A c.4273A>T c.7549-80T>A	**p.Ile1425Phe**	28	3 44	intron A1 intron	het het het	15	M	O+	12	15	8	47	0.53	13	
0118	01036	c.4196G>A	p.Arg1399His †	28		A1	het	60	F	O+	7	16	24	17	1.50	20	1.48
0128	01052*	c.5837G>T	p.Cys1946Phe	34		D4	het	51	F	O-	16	30	18	68	0.60	18	
0132	01066	c.4244G>A	p.Gly1415Asp	28		A1	het	20	F	O+	17	16	4	26	0.25	11	
	01067	c.4244G>A	p.Gly1415Asp	28		A1	het	53	M	O+	8	22	4	35	0.18	16	
0216	02045	c.3961T>A	p.Tyr1321Asn	28		A1	het	63	F	A+	12	71	18.2	97	0.26	54	
0230	02079	c.4133C>T	p.Ser1378Phe	28		A1	het	50	M	A-	3	46	8.6	68	0.19	42	
	02080	c.4133C>T	p.Ser1378Phe	28		A1	het	15	M	ND	2	40	8.6	72	0.22	32	
0332	03023	c.4225G>T c.6187C>T	**p.Val1409Phe** p.Pro2063Ser	28 36		A1 D4	het het	48	F	A+	12	12	4.4	21	0.37	7.7	
	03024	c.4196G>A c.4225G>T	p.Arg1399His † **p.Val1409Phe**	28 28		A1 A1	het het	22	M	ND	5	25	7.2	26	0.29	12	0.1
	03025	c.4225G>T	**p.Val1409Phe**	28		A1	het	77	F	O+	8	109	17	65	0.16	42	
0382	03045*	c.3788C>T	p.Ser1263Leu	28		D3	het	24	M	O-	3	78	48	20	0.62	62	
0512	05013	c.4196G>A	p.Arg1399His †	28		A1	het	41	F	O-	3	50	47.6	78	0.95	39	6.14
0705	07007	c.4196G>A c.6847T>C	p.Arg1399His † p.Cys2283Arg	28 39		A1 C1	het het	27	F	O+	10	32	39	97	1.22	24	0
1213	12020	c.4196G>A c.7730-177G>T	p.Arg1399His †	28	 45	A1 intron	het het	28	M	O-	0	84	55	67	0.65	53	4.59
1216	12023	c.4183C>T c.533-43A>C	p.Arg1395Trp ‡	28	 5	A1 intron	het het	71	M	ND	8	272	186	213	0.68	178	
2307	23010*	c.1110-26T>A c.1156+27C>T c.7773C>T	 c.7773C>T (p.=)	 47	9 10	intron intron C5	het het het	39	F	O-	4	44	29	70	0.66	36	
2601	26001	c.4121G>T	p.Arg1374Leu	28		A1	het	37	M	B+	9	12	5	38	0.42	11	
	26002	c.4121G>T c.6068C>T	p.Arg1374Leu p.Thr2023Met	28 36		A1 D4	het het	7	M	B+	-1	30	12.2	62	0.41	26	
	26003	c.4121G>T	p.Arg1374Leu	28		A1	het	40	F	B+	13	15	6.8	37	0.45	14	
	26004	c.4121G>T	p.Arg1374Leu	28		A1	het	13	M	O+	10	8.3	5	38	0.60	9	
2706	27009	c.4145T>C	**p.Leu1382Pro**	28		A1	het	39	M	O-	2	5.4	4	15	0.74	5.6	
	27010	c.4145T>C c.4196G>A	**p.Leu1382Pro** p.Arg1399His †	28 28		A1 A1	het het	44	M	O-	1	6.8	4	18	0.59	4.3	ND
	27011	c.4145T>C	**p.Leu1382Pro**	28		A1	het	73	M	O-	18	32	14	70	0.44	26	
2712	27022*	c.5191T>A c.-1873A>G	p.Ser1731Thr §	30		A3 upstream	het het	19	F	O+	1	38	35.5	65	0.93	41	72.2
2901	29001	c.4244G>A	p.Gly1415Asp	28		A1	het	35	M	A-	10	29	10.4	46	0.36	22	
3010	30027	c.5336G>T	p.Arg1779Leu	31		A3	het	38	F	O-	11	36	16	34	0.44	20	
	30028	c.5336G>T	p.Arg1779Leu	31		A3	het	15	F	A-	0	28	13.6	19	0.49	16	

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB	VWF:CB(VI)
3608	36010	c.4157C>A	p.Ala1386Asp	28		A1	het	40	M	O-	1	30	6	33	0.20	21	
		c.6433C>G	p.Pro2145Ala	37		D4	het										
3803	38003	c.4313T>G	p.Phe1438Cys	28		A1	het	15	F	O+	21	28	18	30	0.64	24	
		c.-2077A>T				upstream	het										
3901	39002	c.4222_4224del	**p.Lys1408del**	28		A1	het	66	F	A+	19	21	9	43	0.43	27	
	39003	c.4222_4224del	**p.Lys1408del**	28		A1	het	39	F	A-	9	25	10	50	0.40	30	
	39007	c.4222_4224del	**p.Lys1408del**	28		A1	het	32	M	A+	7	17	5.4	33	0.32	17	
4402	44002	c.4244G>A	p.Gly1415Asp	28		A1	het	43	F	ND	5	13	5	20	0.38	11	
	44003	c.2561G>A	p.Arg854Gln	20		D'	het	6	F	ND	4	16	5	11	0.31	11	
		c.4244G>A	p.Gly1415Asp	28		A1	het										
4406	44007	c.4225G>T	**p.Val1409Phe**	28		A1	het	51	F	O-	3	28	13	37	0.46	21	
	44008	c.4225G>T	**p.Val1409Phe**	28		A1	het	56	M	O+	7	28	6.2	23	0.22	19	
	44009	c.2771G>A	p.Arg924Gln	21		D3	het	22	M	O+	9	46	10.4	60	0.23	32	
		c.4225G>T	**p.Val1409Phe**	28		A1	het										
	44010	c.2771G>A	p.Arg924Gln	21		D3	het	19	M	O+	2	33	7.4	33	0.22	21	
		c.4225G>T	**p.Val1409Phe**	28		A1	het										

In bold, mutations previously described in type 2M. ND indicates not determined. *Patients with a discrepant phenotype-genotype correlation. † Mutations causing a defective binding to collagen type VI. ‡ Mutation causing a defective binding to collagen type IV. § Mutation causing a defective binding to collagen types I,III.

Table S9. Genotype-phenotype correlation in type 2N patients

Family	Patient	Phenotype	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	VWF:Ag	VWF:RCo	FVIII:C	VWF/FVIIIB
0107	01011	2N	c.2446C>T	**p.Arg816Trp**	19		D'	hom	56	F	A+	8	59	86	4.5	0.00
	01016	2N (Carrier)	c.2446C>T	**p.Arg816Trp**	19		D'	het	25	F	A-	1	74	95	124	0.56
0124	01044	2N (Carrier)	c.2561G>A	**p.Arg854Gln**	20		D'	het	40	M	A+	0	52	82	46	0.65
			c.6187C>T	p.Pro2063Ser	36		D4	het								
	01058	2N	c.2561G>A	**p.Arg854Gln**	20		D'	het	31	M	O+	9	42	40	12	0.02
			c.970C>T	p.Arg324Ter	8		null allele	het								
0131	01063	2N (Carrier)	c.2561G>A	**p.Arg854Gln**	20		D'	het	42	F	O+	3	62	61	46	0.65
0202	02002	2N	c.2561G>A	**p.Arg854Gln**	20		D'	het	44	F	O+	8	33	33.6	14	0.05
			c.2685G>C	**p.Gln895His**	20		D3	het								
	02003	2N	c.2561G>A	**p.Arg854Gln**	20		D'	het	45	F	O+	6	32	31.5	16	0.06
			c.2685G>C	**p.Gln895His**	20		D3	het								
	02005	2N (Carrier)	c.2561G>A	**p.Arg854Gln**	20		D'	het	68	F	B+	0	128	149	110	0.52
	02006	2N (Carrier)	c.2561G>A	**p.Arg854Gln**	20		D'	het	12	F	O+	0	67	71	114	0.49
	02008	2N (Carrier)	c.2561G>A	**p.Arg854Gln**	20		D'	het	7	F	O-	0	73	101	71	0.66
			c.7438-31T>C			43	intronic	het								
			c.7936C>A	p.Pro2646Thr	48		C5	het								
0211	02031	2N	c.2561G>A	**p.Arg854Gln**	20		D'	het	41	M	O+	6	44	56	24	0.11
			c.7672_7676del	p.Pro2558GlyfsTer7	45		null allele	het								
	02032	2N (Carrier)	c.1654G>A	p.Ala552Thr	14		D2	het	13	F	ND	1	82	84.8	108	0.63
			c.2561G>A	**p.Arg854Gln**	20		D'	het								
			c.3538+20G>A			26	intronic	het								
	02033	2N (Carrier)	c.2546+97C>G			19	intronic	het	9	M	ND	0	65	80	111	0.62
			c.2561G>A	**p.Arg854Gln**	20		D'	het								
0229	02075	2N (Carrier)	c.2561G>A	**p.Arg854Gln**	20		D'	het	52	M	O+	5	55	54.4	54	0.7
			c.3144C>T	c.3144C>T (p.=)	24		D3	het								
			c.546G>A	c.546G>A (p.=)	6		D1	het								
	02076	2N (Carrier)	c.2561G>A	**p.Arg854Gln**	20		D'	het	18	M	ND	0	34	34.4	54	0.57
			c.3144C>T	c.3144C>T (p.=)	24		D3	het								
0319	03014	2N	c.2446C>T	**p.Arg816Trp**	19		D'	hom	67	F	A+	11	274	171	5.1	0.00
0513	05014	2N	c.2446C>T	**p.Arg816Trp**	19		D'	hom	39	F	A+	11	51	48.8	5.2	0.00
1406	14006	2N	c.2561G>A	**p.Arg854Gln**	20		D'	hom	17	F	ND	16	102	110	3	0.1
2105	21008	2N	c.2446C>T	**p.Arg816Trp**	19		D'	hom	44	M	O+	7	114	93	7.5	0.00
3604	36006	2N (Carrier)	c.1293+86C>T			11	intronic	het	4	M	A+	0	75	60	57	0.44
			c.2561G>A	**p.Arg854Gln**	20		D'	het								

In bold, mutations previously described in type 2N. ND indicates not determined.

Table S10. Patients with uncertain classification

Family	Patient	HGVSc	HGVSp	Exon	Intron	Domain	Genotype	Age	Gender	ABO	BS	Multimeric pattern	VWF:Ag	VWF:RCo	FVIII:C	VWF:RCo/VWF:Ag	VWF:CB	Potential classification*
0109	01017	c.220+100G>A			3	intronic	het	39	F	O+	3	NORMAL	36	49	101	1.36	46	-
		c.7549-80T>A			44	intronic	het											
	01074	c.220+100G>A			3	intronic	het	17	F	O+	2	NORMAL	39	41	96	1.05	41	-
		c.7549-80T>A			44	intronic	het											
0901	09001	c.4775T>A	p.Val1592Asp	28		A2	het	9	M	O	0	NORMAL	104	90	70	0.87	82	-
		c.7888-65C>A			47	intronic	het											
1003	10004	c.3829G>C	p.Asp1277His	28		A1	het	38	F	O+	3	ND	46	19.8	43	0.43	25	2A
1203	12006	c.1110-26T>A			9	intronic	het	29	F	O+	0	NORMAL	42	38.5	55	0.92	33	-
		c.1646T>C	p.Phe549Ser	14		D2	het											
1209	12016	c.2900G>A	p.Gly967Asp	22		D3	het	61	F	O+	0	NORMAL	75	62	48	0.83	56	
1901	19001	c.4627T>C	p.Ser1543Pro	28		A2	het	31	M	O-	9	↓HMWM	34	9.1	55	0.27	17	2A
	19002	c.4627T>C	p.Ser1543Pro	28		A2	het	55	F	O+	17	↓HMWM	25	9.3	51	0.37	14	2A
1902	19003	c.4574T>G	p.Ile1525Ser	28		A2	het	34	F	O+	3	↓HMWM	29	12	41	0.41	17	2A
	19004	c.4574T>G	p.Ile1525Ser	28		A2	het	27	F	O+	0	↓HMWM	46	24	51	0.52	25	2A
2302	23004	c.7493C>A	p.Ala2498Asp	44		C4	het	35	F	O+	3	NORMAL	64	54.5	78	0.85	53	-
2701	27002	c.7025G>A	p.Arg2342His	41		C2	het	32	F	ND	-2	NORMAL	46	42.7	72	0.93	48	-
3002	30018	c.2821-88A>G			21	intronic	het	19	M	A-	5	↓↓HMWM	15	5	15	0.33	4.6	-
		c.-2077A>T				upstream	het											
3905	39015	c.-138A>G		1		upstream	het	44	F	ND	-1	NORMAL	98	97	88	0.99	96	-
		c.3539-35G>C			26	intronic	het											

ND indicates not determined. *Potential classification on the basis of molecular results, even though we preferred achieve further clinical and phenotypic evidence to establish a definite classification. In this line, five patients were potentially classified as 2A since the mutation found in each case was located outside the 16 residues related to type 2B.

Table S11. Missense candidate mutations in silico analysis

HGVSc	HGVSp	Exon	Domain	PolyPhen	Sift	TASTERprediction	Mut Assesor	Provean Prediction (cutoff=-25)	*in silicio* global score	#Patients	#Families	Potencial ancestral origin
c.55G>A	p.Gly19Arg	2	SigNDI peptide	possibly_damaging (0.884)	deleterious (0)	disease causing	medium	Neutral	**4**	1	1	-
c.250C>T	**p.Leu84Phe**	4	D1	probably_damaging (0.985)	tolerated (0.15)	disease causing	low	Neutral	2	1	1	-
c.425G>A	**p.Gly142Asp**	5	D1	possibly_damaging (0.818)	tolerated (0.28)	polymorphism	medium	Neutral	2	1	1	-
c.449T>C	p.Leu150Pro	5	D1	probably_damaging (0.972)	deleterious (0)	disease causing	medium	Deleterious	**5**	1	1	-
c.478G>A	**p.Gly160Arg**	5	D1	possibly_damaging (0.854)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.818G>C	**p.Arg273Pro**	7	D1	probably_damaging (0.963)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.1001G>A	**p.Gly334Glu**	9	D1	probably_damaging (0.998)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.1109G>A	**p.Cys370Tyr**	9	D1	probably_damaging (0.933)	deleterious (0.01)	disease causing	high	Deleterious	**5**	6	1	-
c.1446C>G	**p.Ile482Met**	13	D2	probably_damaging (0.998)	deleterious (0)	disease causing	medium	Neutral	**4**	3	3	NO
c.1472G>C	**p.Arg491Pro**	13	D2	probably_damaging (0.936)	deleterious (0.01)	disease causing	medium	Deleterious	**5**	2	1	-
c.1625C>G	p.Ala542Gly	14	D2	possibly_damaging (0.539)	deleterious (0)	polymorphism	low	Deleterious	**3**	5	5	NO
c.1646T>C	**p.Phe549Ser**	14	D2	probably_damaging (0.998)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.1654G>A	**p.Ala552Thr**	14	D2	benign (0.363)	tolerated (0.3)	polymorphism	low	Neutral	0	1	1	-
c.1721C>G	**p.Pro574Arg**	14	D2	probably_damaging (1)	deleterious (0.01)	disease causing	medium	Deleterious	**5**	1	1	-
c.1847C>T	**p.Ser616Leu**	15	D2	probably_damaging (0.973)	deleterious (0.04)		medium	Deleterious	**5**	1	1	-
c.2119T>C	**p.Cys707Arg**	16	D2	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.2289G>C	**p.Arg763Ser**	18	D2	probably_damaging (0.994)	deleterious (0)	disease causing	medium	Deleterious	**5**	3	2	YES
c.2446C>T	p.Arg816Trp	19	D'	probably_damaging (1)	deleterious (0)	disease causing	medium	Deleterious	**5**	6	5	NO
c.2561G>A	p.Arg854Gln	20	D'	probably_damaging (0.996)	deleterious (0.01)	disease causing	low	Deleterious	**4**	19	11	NO
c.2685G>C	p.Gln895His	20	D3	probably_damaging (0.987)	deleterious (0)	disease causing	medium	Deleterious	**5**	7	2	YES
c.2771G>A	p.Arg924Gln	21	D3	possibly_damaging (0.635)	tolerated (0.36)	polymorphism	low	Neutral	1	15	13	NO
c.2878C>T	p.Arg960Trp	22	D3	probably_damaging (0.997)	deleterious (0.03)	disease causing	medium	Neutral	**4**	2	2	NO
c.2900G>A	p.Gly967Asp	22	D3	probably_damaging (0.991)	tolerated (0.21)	disease causing	low	Neutral	2	1	1	-
c.2900G>T	**p.Gly967Val**	22	D3	probably_damaging (0.998)	deleterious (0)	disease causing	medium	Deleterious	**5**	1	1	-
c.2926C>T	**p.Arg976Cys**	22	D3	probably_damaging (0.939)	deleterious (0.01)	polymorphism	medium	Deleterious	**4**	2	1	-
c.2944G>C	**p.Val982Leu**	22	D3	probably_damaging (0.999)	tolerated (0.06)	disease causing	low	Neutral	2	1	1	-
c.3319T>C	**p.Tyr1107His**	25	D3	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.3363G>T	**p.Arg1121Ser**	25	D3	probably_damaging (0.999)	deleterious (0)	disease causing	high	Deleterious	**5**	4	2	YES
c.3433C>T	**p.Arg1145Cys**	26	D3	probably_damaging (1)	deleterious (0)	disease causing	medium	Deleterious	**5**	1	1	-
c.3445T>C	p.Cys1149Arg	26	D3	probably_damaging (0.999)	deleterious (0)	disease causing	high	Deleterious	**5**	11	4	YES
c.3467C>T	p.Thr1156Met	26	D3	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	7	4	NO
c.3485_3486delinsTG*	p.Pro1162Leu	26	D3	probably_damaging (1)	deleterious (0.01)		medium	Deleterious	**5**	7	7	NO
c.3614G>A	p.Arg1205His	27	D3	probably_damaging (0.998)	tolerated (0.16)	disease causing	medium	Neutral	**3**	35	15	NO
c.3692A>G	p.Asn1231Ser	28	D3	benign (0.083)	tolerated (0.26)	disease causing	low	Neutral	1	2	2	NO
c.3788C>T	**p.Ser1263Leu**	28	D3	benign (0.005)	tolerated (0.66)	polymorphism	low	Neutral	0	1	1	-
c.3797C>A	p.Pro1266Gln	28	D3	probably_damaging (0.999)	deleterious (0.05)	polymorphism	medium	Neutral	2	2	1	-
c.3797C>T	p.Pro1266Leu	28	D3	probably_damaging (0.995)	tolerated (1)	disease causing	medium	Neutral	**3**	4	3	YES
c.3815G>C	p.Cys1272Ser	28	A1	probably_damaging (1)	deleterious (0)	disease causing	medium	Deleterious	**5**	1	1	-
c.3829G>C	**p.Asp1277His**	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.3835G>A	p.Val1279Ile	28	A1	possibly_damaging (0.725)	tolerated (0.07)	disease causing	high	Neutral	**3**	3	2	NO
c.3916C>T	p.Arg1306Trp	28	A1	probably_damaging (0.998)	deleterious (0)	disease causing	high	Neutral	**4**	11	7	NO
c.3917G>A	p.Arg1306Gln	28	A1	possibly_damaging (0.874)	tolerated (0.12)	disease causing	medium	Neutral	**3**	2	2	NO
c.3920T>C	p.Leu1307Pro	28	A1	probably_damaging (1)	deleterious (0.01)	disease causing	high	Neutral	**4**	3	1	-
c.3922C>T	p.Arg1308Cys	28	A1	probably_damaging (0.925)	deleterious (0.02)	disease causing	low	Neutral	**3**	12	7	NO

HGVSc	HGVSp	Exon	Domain	PolyPhen	Sift	TASTERprediction	Mut Assesor	Provean Prediction (cutoff=-25)	*in silicio* global score	#Patients	#Families	Potencial ancestral origin
c.3943C>T	p.Arg1315Cys	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Neutral	**4**	7	4	NO
c.3944G>T	p.Arg1315Leu	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.3946G>A	p.Val1316Met	28	A1	probably_damaging (1)	deleterious (0.01)	disease causing	high	Neutral	**4**	5	4	NO
c.3961T>A	**p.Tyr1321Asn**	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.4010C>T	p.Pro1337Leu	28	A1	probably_damaging (0.998)	tolerated (0.72)	disease causing	medium	Neutral	**3**	1	1	-
c.4120C>T	p.Arg1374Cys	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Neutral	**4**	22	6	NO
c.4121G>A	p.Arg1374His	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Neutral	**4**	30	7	NO
c.4121G>T	p.Arg1374Leu	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Neutral	**4**	4	1	-
c.4133C>T	p.Ser1378Phe	28	A1	probably_damaging (0.982)	deleterious (0.03)	polymorphism	medium	Deleterious	**4**	2	1	-
c.4145T>C	**p.Leu1382Pro**	28	A1	probably_damaging (1)	deleterious (0)	disease causing	high	Neutral	**4**	3	1	-
c.4157C>A	**p.Ala1386Asp**	28	A1	probably_damaging (1)	deleterious (0)	disease causing	medium	Neutral	**4**	1	1	-
c.4183C>T	**p.Arg1395Trp**	28	A1	probably_damaging (0.981)	deleterious (0)	polymorphism	high	Neutral	**3**	1	1	-
c.4196G>A	p.Arg1399His	28	A1	possibly_damaging (0.549)	deleterious (0.05)	disease causing	medium	Neutral	**4**	8	8	NO
c.4225G>T	**p.Val1409Phe**	28	A1	probably_damaging (0.998)	deleterious (0)	disease causing	high	Neutral	**4**	7	2	YES
c.4238C>T	p.Pro1413Leu	28	A1	probably_damaging (0.989)	tolerated (0.11)	disease causing	medium	Deleterious	**4**	1	1	-
c.4241T>G	**p.Val1414Gly**	28	A1	probably_damaging (0.996)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.4244G>A	p.Gly1415Asp	28	A1	probably_damaging (1)	deleterious (0.02)	disease causing	high	Neutral	**4**	5	3	NO
c.4255C>A	**p.His1419Asn**	28	A1	possibly_damaging (0.884)	tolerated (0.73)	polymorphism	high	Neutral	2	2	1	-
c.4257T>G	**p.His1419Gln**	28	A1	possibly_damaging (0.819)	tolerated (0.22)	polymorphism	medium	Neutral	2	1	1	-
c.4273A>T	p.Ile1425Phe	28	A1	probably_damaging (0.996)	deleterious (0.01)	disease causing	high	Deleterious	**5**	1	1	-
c.4313T>G	**p.Phe1438Cys**	28	A1	probably_damaging (0.958)	deleterious (0.01)	disease causing	high	Neutral	**4**	1	1	-
c.4499C>A	p.Ala1500Glu	28	A2	probably_damaging (0.934)	tolerated (0.17)	polymorphism	medium	Neutral	2	1	1	-
c.4517C>T	p.Ser1506Leu	28	A2	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	10	7	NO
c.4574T>G	**p.Ile1525Ser**	28	A2	probably_damaging (1)	deleterious (0)	disease causing	medium	Neutral	**4**	2	1	-
c.4580G>A	p.Arg1527Gln	28	A2	possibly_damaging (0.784)	tolerated (0.5)	polymorphism	low	Neutral	1	1	1	-
c.4622A>T	**p.Gln1541Leu**	28	A2	probably_damaging (1)	deleterious (0.02)	disease causing	high	Deleterious	**5**	1	1	-
c.4627T>C	**p.Ser1543Pro**	28	A2	probably_damaging (0.999)	deleterious (0.01)	disease causing	high	Neutral	**4**	2	1	-
c.4735G>A	p.Gly1579Arg	28	A2	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	1	1	-
c.4739T>C	p.Leu1580Pro	28	A2	probably_damaging (0.99)	tolerated (0.21)	polymorphism	high	Neutral	2	1	1	-
c.4747C>T	p.Arg1583Trp	28	A2	possibly_damaging (0.88)	deleterious (0)	polymorphism	low	Neutral	2	1	1	-
c.4751A>G	p.Tyr1584Cys	28	A2	probably_damaging (0.981)	tolerated (0.09)	polymorphism	high	Neutral	2	8	8	NO
c.4775T>A	**p.Val1592Asp**	28	A2	benign (0.201)	tolerated (0.56)	polymorphism	neutral	Neutral	1	1	1	-
c.4789C>G	p.Arg1597Gly	28	A2	probably_damaging (0.999)	tolerated (0.16)	polymorphism	high	Neutral	2	2	1	-
c.4789C>T	p.Arg1597Trp	28	A2	probably_damaging (1)	deleterious (0)	polymorphism	high	Deleterious	**4**	9	5	NO
c.4790G>A	p.Arg1597Gln	28	A2	probably_damaging (0.984)	tolerated (0.05)	polymorphism	high	Neutral	2	4	1	-
c.4825G>A	p.Gly1609Arg	28	A2	probably_damaging (1)	tolerated (0.23)	polymorphism	high	Neutral	2	4	3	NO
c.4840G>A	**p.Asp1614Asn**	28	A2	probably_damaging (1)	deleterious (0.02)	disease causing	high	Neutral	**4**	3	1	-
c.4883T>C	p.Ile1628Thr	28	A2	probably_damaging (0.998)	deleterious (0.01)	disease causing	high	Neutral	**4**	11	6	NO
c.4885G>A	p.Gly1629Arg	28	A2	probably_damaging (1)	deleterious (0.03)	disease causing	high	Deleterious	**5**	1	1	-
c.4889T>G	**p.Val1630Gly**	28	A2	probably_damaging (1)	deleterious (0)	disease causing	high	Deleterious	**5**	2	1	-
c.4892G>A	**p.Gly1631Asp**	28	A2	probably_damaging (1)	deleterious (0.03)	disease causing	high	Deleterious	**5**	1	1	-
c.4943C>G	p.Pro1648Arg	28	A2	probably_damaging (1)	deleterious (0.01)	disease causing	high	Deleterious	**5**	1	1	-
c.4960T>C	p.Phe1654Leu	28	A2	probably_damaging (0.999)	tolerated (0.06)	polymorphism	high	Neutral	2	4	2	YES
c.5191T>A	p.Ser1731Thr	30	A3	probably_damaging (0.996)	deleterious (0.02)	disease causing	medium	Neutral	**4**	2	2	NO
c.5198T>C	**p.Leu1733Pro**	30	A3	probably_damaging (0.995)	deleterious (0)	disease causing	medium	Neutral	**4**	12	6	NO
c.5278G>A	p.Val1760Ile	30	A3	benign (0.043)	tolerated (0.77)	polymorphism	low	Neutral	0	3	2	NO
c.5311G>A	**p.Gly1771Arg**	30	A3	possibly_damaging (0.774)	deleterious (0.03)	disease causing	medium	Deleterious	**4**	1	1	-

HGVSc	HGVSp	Exon	Domain	PolyPhen	Sift	TASTERprediction	Mut Assesor	Provean Prediction (cutoff=-25)	*in silicio* global score	#Patients	#Families	Potencial ancestral origin
c.5336G>T	**p.Arg1779Leu**	31	A3	possibly_damaging (0.609)	tolerated (0.11)	disease causing	medium	Deleterious	**4**	3	2	YES
c.5368C>T	**p.Pro1790Ser**	31	A3	benign (0.045)	tolerated (0.07)	disease causing	medium	Neutral	2	1	1	-
c.5471C>A	p.Pro1824His	32	A3	probably_damaging (0.955)	deleterious (0)	disease causing	medium	Deleterious	**5**	4	1	-
c.5695T>C	**p.Cys1899Arg**	34	D4	benign (0.358)	deleterious (0)	disease causing	medium	Deleterious	**4**	2	1	-
c.5837G>T	**p.Cys1946Phe**	34	D4	benign (0.214)	deleterious (0)	disease causing	medium	Deleterious	**4**	2	2	YES
c.6068C>T	**p.Thr2023Met**	36	D4	benign (0.032)	tolerated (0.23)	polymorphism	low	Neutral	0	1	1	-
c.6187C>T	p.Pro2063Ser	36	D4	benign (0.372)	deleterious (0.02)	disease causing	medium	Deleterious	**4**	28	18	NO
c.6197A>G	p.Asn2066Ser	36	D4	possibly_damaging (0.534)	deleterious (0.03)	disease causing	medium	Neutral	**4**	8	7	NO
c.6248G>A	**p.Gly2083Asp**	36	D4	probably_damaging (0.986)	deleterious (0)	disease causing	medium	Deleterious	**5**	5	2	YES
c.6254G>T	**p.Cys2085Phe**	36	D4	probably_damaging (0.978)	deleterious (0)	disease causing	medium	Deleterious	**5**	1	1	-
c.6352C>T	**p.Arg2118Trp**	37	D4	benign (0.199)	deleterious (0.01)	polymorphism	neutral	Neutral	2	1	1	-
c.6433C>G	**p.Pro2145Ala**	37	D4	benign (0)	tolerated (0.74)	polymorphism	neutral	Neutral	1	1	1	-
c.6433C>T	p.Pro2145Ser	37	D4	benign (0)	tolerated (0.61)	polymorphism	neutral	Neutral	1	1	1	-
c.6536C>T	p.Ser2179Phe	37	D4	possibly_damaging (0.876)	deleterious (0)	disease causing	medium	Neutral	**4**	2	1	-
c.6847T>C	**p.Cys2283Arg**	39	C1	benign (0.326)	deleterious (0)	disease causing	medium	Deleterious	**4**	1	1	-
c.6911G>A	p.Cys2304Tyr	40	C1	probably_damaging (0.936)	deleterious (0)	disease causing	medium	Deleterious	**5**	1	1	-
c.6989T>G	**p.Val2330Gly**	41	C1	benign (0.001)	tolerated (0.23)	polymorphism	medium	Neutral	1	3	1	-
c.7025G>A	**p.Arg2342His**	41	C2	benign (0.001)	tolerated (0.23)	polymorphism	neutral	Neutral	1	1	1	-
c.7070A>T	**p.Asn2357Ile**	41	C2	possibly_damaging (0.533)	tolerated (0.67)	polymorphism	low	Neutral	0	1	1	-
c.7118C>T	**p.Pro2373Leu**	42	C2	possibly_damaging (0.636)	tolerated (0.21)	disease causing	medium	Deleterious	**4**	1	1	-
c.7393G>A	**p.Val2465Met**	43	C3	probably_damaging (0.96)	tolerated (0.16)	disease causing	low	Neutral	2	2	2	YES
c.7471T>C	p.Cys2491Arg	44	C3	possibly_damaging (0.876)	deleterious (0)	disease causing	medium	Deleterious	**5**	2	1	-
c.7493C>A	**p.Ala2498Asp**	44	C4	possibly_damaging (0.795)	tolerated (0.1)	disease causing	medium	Neutral	**3**	1	1	-
c.7583G>A	**p.Cys2528Tyr**	45	C4	possibly_damaging (0.49)	deleterious (0)	disease causing	medium	Deleterious	**5**	1	1	-
c.7936C>A	**p.Pro2646Thr**	48	C5	benign (0.235)	deleterious (0.01)	disease causing	medium	Neutral	**3**	1	1	-
c.7985A>G	**p.Lys2662Arg**	48	C6	benign (0.301)	tolerated (0.09)	disease causing	medium	Neutral	2	1	1	-
c.8318G>C	p.Cys2773Ser	52	CK	benign (0.004)	deleterious (0)	disease causing	medium	Deleterious	**4**	1	1	-
c.8366C>G	**p.Thr2789Ser**	52	CK	benign (0.054)	tolerated (0.53)	disease causing	medium	Neutral	2	1	1	-

Mutations In bold were not described previously in EAHAD-VWD-LOVD. The *in silico* global score is based on the number of *in silico* algorithms that predicted a deleterious effect. In silico global score >2 are highlighted in bold.

*The c.3485_3486delinsTG (p.Pro1162Leu) was also aNDlysed by localSpliceEffect program with a prediction of cryptic acceptor strongly activated site.

Table S12. Nonsense mutations

HGVSc	HGVSp	Exon	Domain	Consequence	#Patients	#Families	Potencial ancestral origin
c.100C>T	p.Arg34Ter	3	D1	stop gained	3	2	NO
c.246_247insT	**p.Ser83Ter**	4	D1	frameshift variant, feature elongation	1	1	-
c.324-2_326dup	**p.Ser110GlufsTer13**	5	D1	frameshift variant, feature elongation	1	1	-
c.375_376delInsC	**p.Tyr126ThrfsTer49**	5	D1	frameshift variant, feature truncation	8	7	NO
c.970C>T	p.Arg324Ter	8	D1	stop gained	5	3	NO
c.1093C>T	p.Arg365Ter	9	D1	stop gained	1	1	-
c.1209C>G	**p.Tyr403Ter**	11	D2	stop gained	1	1	-
c.1992dupC	**p.Cys665LeufsTer13**	16	D2	frameshift variant, feature elongation	1	1	-
c.(1945+1_1946-1)_(7437+1_7438-1)del	**-**	16-43	D2-C3	frameshift variant	1	1	-
c.2435delC	p.Pro812ArgfsTer31	18	D'	frameshift variant, feature truncation	1	1	-
c.2518G>T	**p.Glu840Ter**	19	D'	stop gained	2	1	-
c.2540dupA	**p.Asn847LysfsTer18**	19	D'	frameshift variant, feature elongation	2	1	-
c.3835_4105conNG_001212.3:g.6566_6836	**p.[(V1279I;Q1311*;I1343V;V1360A;F1369I)]**	28	A1	gene conversion	1	1	-
c.3931C>T	p.Gln1311Ter	28	A1	stop gained	8	6	NO
c.4162C>T	**p.Gln1388Ter**	28	A1	stop gained	2	2	YES
c.4975C>T	p.Arg1659Ter	28	A2	stop gained	2	2	NO
c.6182delT	p.Phe2061SerfsTer38	36	D4	frameshift variant, feature truncation	1	1	-
c.6699_6702dup	**p.Cys2235ArgfsTer8**	38	D4	frameshift variant, feature elongation	1	1	-
c.7300C>T	p.Arg2434Ter	43	C3	stop gained	1	1	-
c.7408C>T	p.Gln2470Ter	43	C3	stop gained	4	4	YES
c.7664_7665insAG	p.Cys2557SerfsTer8	45	C4	frameshift variant, feature elongation	1	1	-
c.7672_7676del	**p.Pro2558GlyfsTer7**	45	C4	frameshift variant, feature truncation	1	1	-
c.8347C>T	**p.Gln2783Ter**	52	CK	stop gained	4	3	NO

In bold, mutations not described previously in EAHAD-VWD-LOVD. Since all are null alleles, suffering potential NMD, domains are only indicated for informative positioNDl purposes.

Tabla S13. In frame candidate mutations

HGVSc	HGVSp	Exon	Domain	Genotype	Combined with		#Patients	#Families	Potencial ancestral origin
c.1157_5620del	p.Gly386_Ser1873del	11-32	D2-A3	het	-	2A	1	1	-
c.1534_1536del	p.Leu512del	14	D2	het	p.Tyr126ThrfsTer49 and p.Asn2066Ser	3	1	1	-
c.3223-7_3236dup	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	25	D3	het	-	1	5	2	YES
c.4222_4224del	p.Lys1408del	28	A1	het	-	2M	3	1	-
c.4678_4680dup	p.Asp1560dup	28	NP	het	-	2A	2	1	-

All mutations are new.

Tabla S14. Splicing candidate mutations

HGVSc	Intron	wtGSScore	varGSScore	wtSSFScore	varSSFScore	wtHSFScore	varHSFScore	HSF score variation (%)	wtNNSScore	varNNSScore	wtMaxEntScore	varMaxEntScore	Combinen HSF&MaxEnt conclusion	#Patients	#Families	Potencial ancestral origin
c.533-2A>G	5	8.48	NP	88.84	NP	96.35	67.4	30.05 WT site broken	0.90	NP	9.11	1.15	Alteration of the WT donor site. most probably affecting splicing.	3	3	YES
c.874+8G>A	7	-	-	-	-	43.86	72.8	+65.98 new aceptor site creation	-	-	-	-	No significant splicing motif alteration detected.	1	1	-
c.1156+2T>C	10	8.55	NP	84.14	82.05	89.49	62.65	29.99 WT site broken	0.95	NP	9.88	2.12	Alteration of the WT donor site. most probably affecting splicing.	2	1	-
c.1533+1G>A	13	14.21	NP	89.61	NP	90.71	63.87	29.59 WT site broken	0.98	NP	10.29	2.10	Alteration of the WT donor site. most probably affecting splicing.	1	1	-
c.3379+1G>A	25	7.25	NP	70.36	NP	81.32	54.48	33.01 WT site broken	0.90	NP	8.56	0.38	Alteration of the WT donor site. most probably affecting splicing.	1	1	-
c.5455+1G>A	31	5.57	NP	73.31	NP	81.38	54.55	32.97 WT site broken	0.89	NP	5.96	-2.21	Alteration of the WT donor site. most probably affecting splicing.	1	1	-
c.6256+2dupT	36	2.02	NP	83.11	61.39	86.69	74.37	18.41 WT site broken	0.99	0.002	8.24	-4.77	Alteration of the WT donor site. most probably affecting splicing.	1	1	-
c.6598+1G>A	37	8.03	NP	84.07	NP	92.33	65.5	29.06 WT site broken	0.99	NP	8.95	0.76	Alteration of the WT donor site. most probably affecting splicing. Activation of an intronic cryptic donor site 3 nt upstream.	3	1	-
c.7082-2A>G	41	3.41	NP	79.01	NP	88.28	59.34	32.78 WT site broken	0.69	NP	6.68	-1.27	Alteration of the WT donor site. most probably affecting splicing.	3	2	NO
c.7730-4C>G	45	13.44	14.55	87.12	87.12	91.73	91.11	-0.68. New site 3 nt upstream	0.99	0.99	11.71	13.00	Activation of an intronic cryptic acceptor site. Potential alteration of splicing.	1	1	-
c.7730-1G>C	45	13.44	NP	87.12	NP	91.73	62.79	31.55 WT site broken	0.99	NP	11.71	3.65	Alteration of the WT donor site. most probably affecting splicing.	6	5	NO
c.8155+3G>C	50	1.81	NP	83.02	77.50	85.46	78.65	-7.97	0.94	0.21	7.83	2.02	Alteration of the WT donor site. most probably affecting splicing.	2	2	YES
c.8254-5T>G	51	15.94	12.23	93.49	89.36	92.50	88.88	-3.91	0.99	0.94	12.47	10.09	No significant splicing motif alteration detected. This mutation has probably no impact on splicing.	2	2	YES

A splice site effect was considered as potentially deleterious when a variation between the NDtive and the mutation score of more than 10% was observed in both algorithms. In bold, mutations not described previously in EAHAD-VWD-LOVD. GS indicates GeneSplicer; SSF, Splicing Sequences Finder; HSF, Human Splicing Finder; NNS, Neural Network Splice; MaxEnt, Maximum Entropy; NP, not predicted; and WT, wild type.

Table S15. Upstream variants

HGVSc	Genotype		Combined with	#Patients	#Families	Comments	Potencial ancestral origin
c.-3151T>G	het	1	p.Thr1156Met	5	3	-	YES
	het	3	p.Thr1156Met and c.7730-1G>C	1	1		
	hom	1	p.Thr1156Met	1	1		
c.-2692C>T	het	2B	p.Arg1308Cys	1	1	-	-
c.-2627C>T	het	1	c.3379+1G>A	1	1	-	NO
		2A/2M	p.Arg1315Cys	1	1		
c.-2077A>T	het	1	p.Arg1205His	2	2	-	NO
		2M	p.Phe1438Cys	1	1		
		UC	None	1	1		
c.-2076A>G	het	2A	p.Arg1374His	1	1	Within the 20 nt close to the second cluster for DNDse I	-
c.-1896C>T	het	2A/2M	p.Arg1315Cys	2	1	-	-
c.-1875G>A	het	2A	p.Arg1145Cys and p.Arg1374His	1	1	Within the 10 nt close to the first DNDse I cluster	YES
			p.Arg1374His	2	2		
c.-1873A>G	het	2M	p.Ser1731Thr	1	1	Within the 10 nt close to the first DNDse I cluster	-
c.-1650G>C*	het	2A	p.Leu1307Pro	1	1	Within the first Pol 2 and DNDse I cluster	-
c.-138A>G	het	UC	None	1	1	Within uORF mRND regulador element, and in Exon enhancer	-

ENCODE project datasets was used to determine the colocalization of these upstream mutations with a well defined regulatory elements (DNDseI Hypersensitivity Clusters, TF motif sequences determined by ChIPseq high throughput sequencing) in the genome. In bold, mutations not described previously in EAHAD-VWD-LOVD. Of interest, except for 3 patients (2 with uncertain classification and 1 type 3), all upstream mutations were combined with another mutation that could explained the phenotypes. uORFs indicates upstream open reading frames defined as a major gene expression regulatory elements; and UC indicates uncertain classification.* This mutation is located into the consensus sequence for the zinc finger transcription factor GATA binding protein 2 (GATA2).

Tabla S16. Synonymous candidate mutations

HGVSc	HGVSp	Exon	Domain	Adjacent nucleotides	phastCons	LocalSpliceEffect	wtHSFScore	varHSFScore	HSF score variation (%)	wtMaxEntScore	varMaxEntScore	HSF&MaxEnt conclusion	#Patients	#Families	Potencial ancestral origin
c.546G>A	**p.Ser182**	6	D1	tcGga>tcAga	0.13	New Acceptor Site	57.31	86.26	50.51 New site	NP	NP	Activation of an exonic cryptic acceptor site. with presence of one or more cryptic branch point (s). Potential alteration of splicing.	3	3	NO
c.2025G>A	**p.Pro675**	16	D2	ccGga>ccAga	1.00	New Acceptor Site	59.55	88.5	48.61 New site	1.26	9.22	Activation of an exonic cryptic acceptor site. with presence of one or more cryptic branch point (s). Potential alteration of splicing.	1	1	-
c.2103C>T	**p.Cys701**	16	D2	tgCgt>tgTgt	0.91		44.46	71.29	60.35 New site	NP	NP	Activation of an exonic cryptic donor site. Potential alteration of splicing.	1	1	-
c.2586G>T	**p.Val862**	20	D'	gtGtg>gtTtg	0.99		70	43.16	38.34 Site broken	NP	NP	Activation of an exonic ESE site. Potential alteration of splicing.	1	1	-
c.3144C>T	p.Asn1048	24	D3	aaCaa>aaTaa	0.51		NP	NP	NP	NP	NP	-	2	1	-
c.3390C>T	p.Cys1130	26	D3	tgCga>tgTga	0.98	New Donor Site	52.69	79.53	50.94 New site	NP	NP	Activation of an exonic cryptic donor site. Potential alteration of splicing.	5	3	NO
c.3426T>C	**p.Cys1142**	26	D3	tgTga>tgCga	0.95		82.16	55.33	32.66 Site broken	6.96	0.79	Activation of an exonic ESE site. Potential alteration of splicing.	7	7	NO
c.3735G>A	**p.Val1245**	28	D3	gtGcc>gtAcc	0.07		71.51	72.67	1.62	0.91	3.07	Activation of an exonic ESE site. Potential alteration of splicing.	1	1	-
c.4146G>T	p.Leu1382	28	A1	ctGct>ctTct	1.00	Cryptic Acceptor Strongly Activated	NP	NP	NP	NP	NP	Creation of an exonic ESS site. Potential alteration of splicing.	8	8	NO
c.4620G>C	**p.Leu1540**	28	A2	ctGca>ctCca	0.99		79.7	80.33	0.79	NP	NP	Alteration of an exonic ESE site. Potential alteration of splicing.	1	1	-
c.5001G>A	**p.Gln1667**	28	A2	caGag>caAag	0.98	Cryptic Donor Strongly Activated	84.98	56.04	34.06 Site broken	4.83	3.91	Creation of an exonic ESS site. Potential alteration of splicing.	1	1	-
c.5277C>T	p.Asp1759	30	A3	gaCgt>gaTgt	0.13	Cryptic Acceptor Strongly Activated	79.7	80.41	0.89	3.72	5.00	Creation of an exonic ESS site. Potential alteration of splicing.	6	4	NO
c.6237A>G	**p.Ser2079**	36	D4	tcAaa>tcGaa	0.00		81.02	80.95	-0.09	NP	NP	Activation of an exonic ESE site. Potential alteration of splicing.	1	1	-
c.7344C>T	p.Cys2448	43	C3	tgCga>tgTga	0.00		40.38	67.21	66.44 New site	NP	NP	Activation of an exonic cryptic donor site. Potential alteration of splicing.	2	2	NO
c.7464C>T	**p.Gly2488**	44	C3	ggCga>ggTga	0.79	New Donor Site	71.05	97.89	37.78	2.72	10.47	No significant splicing motif alteration detected. This mutation has probably no impact on splicing.	1	1	-
c.7773C>T	**p.Pro2591**	47	C5	ccCgg>ccTgg	0.09		69.47	69.05	-0.6	7.77	6.90	Creation of an exonic ESS site. Potential alteration of splicing.	1	1	-

In bold, mutations not described previously in EAHAD-VWD-LOVD. NP indicates not predicted; and HSF, Human Splicing Finder.

Tabla S17. Intronic variants

HGVSc	Intron	LocalSpliceEffect	#Patients	#Families
c.55+69A>G	2		1	1
c.220+100G>A	3		3	1
c.221-70G>T	3		1	1
c.533-43A>C	5		3	3
c.998-46C>T	8		1	1
c.1110-26T>A	9		8	6
c.1156+27C>T	10		2	2
c.1156+42C>T	10		1	1
c.1293+86C>T	11	New Donor Site	1	1
c.1293+109T>C	11		1	1
c.1432+166C>A	12		1	1
c.1533+15G>A	13	New Acceptor Site	7	2
c.1946-17_1946-16insTTT	15		1	1
c.1946-16_1946-15insCTC	15		1	1
c.2546+55G>T	19	Cryptic Acceptor Strongly Activated	1	1
c.2546+97C>G	19		1	1
c.2546+132G>A	19		2	2
c.2685+147G>A	20		1	1
c.2821-148G>A	21		1	1
c.2821-123A>C	21		3	3
c.2821-88A>G	21		2	2
c.2821-78A>C	21		1	1
c.3379+27A>T	25		1	1
c.3538+20G>A	26	Cryptic Donor Strongly Activated	1	1
c.3539-35G>C	26		1	1
c.5170+10C>T	29	New Donor Site	8	7
c.5312-138G>A	30		1	1

HGVSc	Intron	LocalSpliceEffect	#Patients	#Families
c.5312-19A>C	30		1	1
c.5456-62A>G	31		1	1
c.5665-36T>C	33		1	1
c.5842+31C>T	34		1	1
c.6798+32G>A	38		1	1
c.6799-47G>A	38		1	1
c.6799-27C>T	38		1	1
c.6976+111G>A	40		2	1
c.7082-13G>C	41		2	1
c.7438-169G>A	43		1	1
c.7438-31T>C	43		1	1
c.7548+22G>A	44	New Acceptor Site	1	1
c.7549-80T>A	44	New Acceptor Site	3	1
c.7730-238G>A	45		1	1
c.7730-177G>T	45	Cryptic Acceptor Strongly Activated	5	4
c.7730-56C>T	45		1	1
c.7771-87G>A	46	Cryptic Donor Strongly Activated	1	1
c.7771-86G>A	46	Cryptic Donor Strongly Activated	1	1
c.7771-82G>A	46	Cryptic Donor Strongly Activated	1	1
c.7771-49G>A	46		1	1
c.7771-40G>A	46		1	1
c.7771-29G>A	46		1	1
c.7888-65C>A	47		2	2
c.8156-42C>T	50		1	1

In bold, variants/putative mutations not described previously in EAHAD-VWD-LOVD.

SUPPLEMENTAL FIGURES

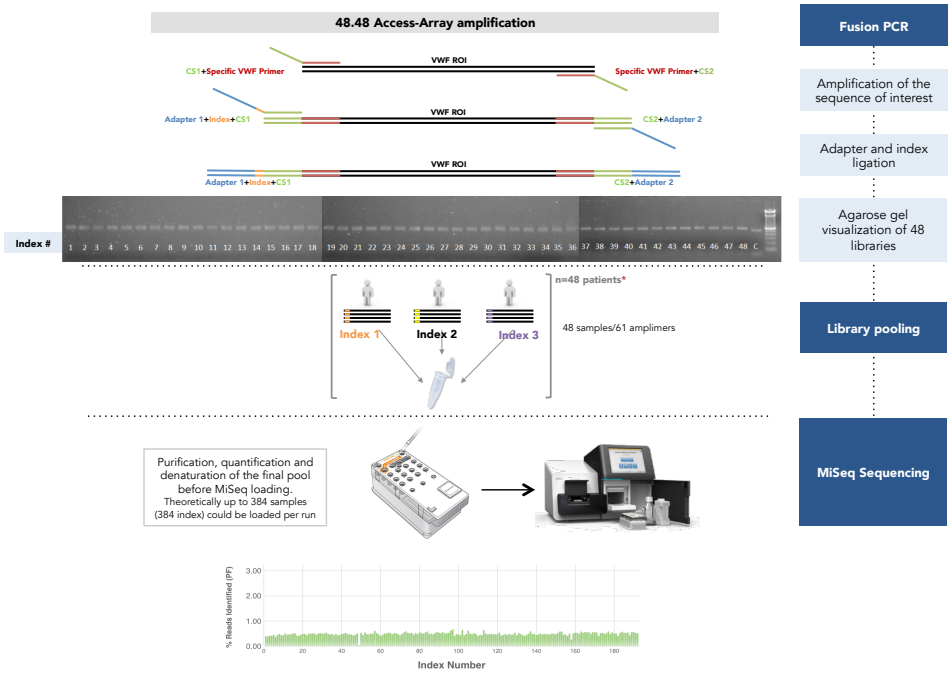


Figure S1. Schematic representation of *VWF* amplification by the Fluidigim Access Array and sequencing on an Illumina MiSeq. The Access Array method amplifies the VWF regions of interest (ROI: exons 1 to 52, intronic flanking regions, and approximately 1,300 bp of the promoter region, resulting in 25,625 bp per patient) by 61 custom primer pairs (red). Each primer pair contains common flanking sequences (CS1 and CS2, green) that permit attachment of a platform-specific barcode index (orange) and sequencing adaptors (blue) in a fusion PCR. The outcome, a pool of 61 amplicons indexed per patient, is known as a library. Each library is visualized on an agarose gel (C is the control size of a library without the adapters, hence the final libraries must be larger than the C band size). As the concentrations of the 48 PCR pools from the same array were highly uniform, a final pool containing a mixture of all samples in the same proportion was created. The final pool of up to 192 samples (4 Access Array) was purified with Agencourt AMPure XP beads (Beckman Coulter Genomics), adjusted to 4 nM, and denatured in 0.2 N NaOH before loading onto a MiSeq Illumina instrument.

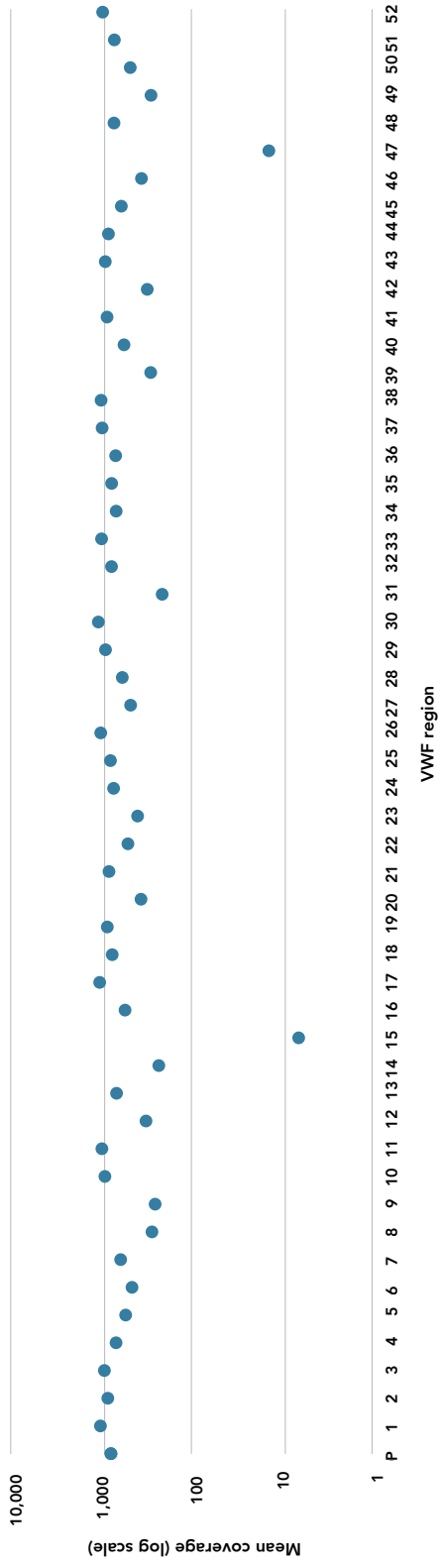


Figure S2. Mean exon coverage graphical representation. Homogeneous output is shown for the promoter and the 52 exons except for exon 47 and exon 15 where amplification failed in a significant number of arrays (undoubtedly due to their high GC content). P indicates promoter region.

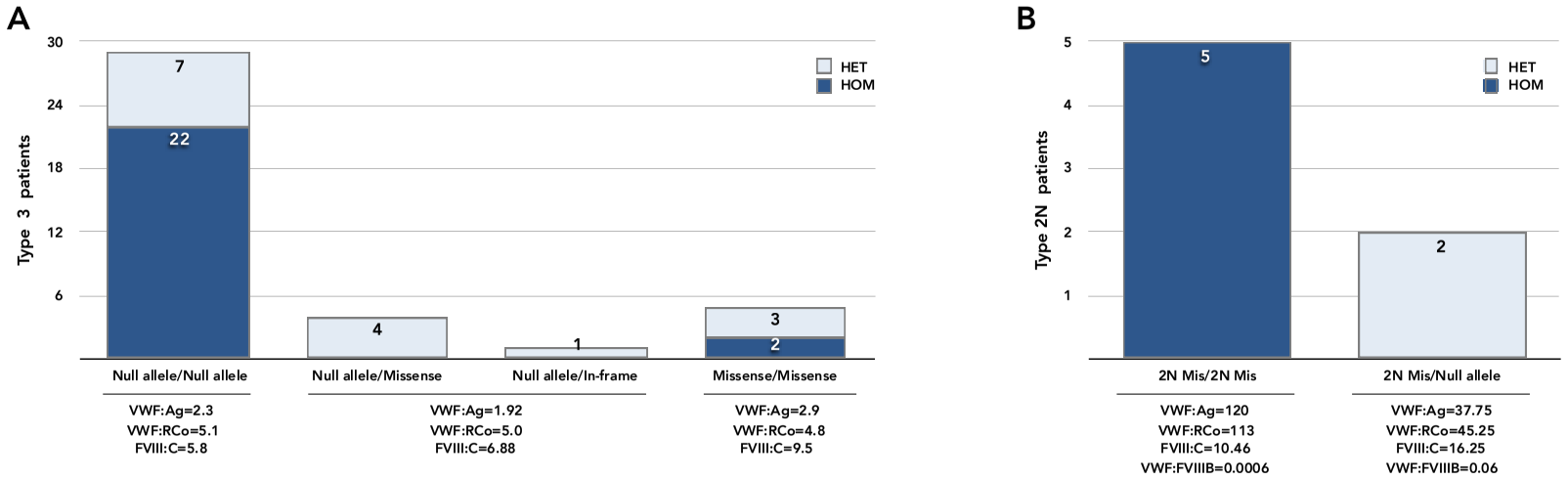


Figure S3. Graphical representation of combined mutations found in recessive VWD subtypes in all phenotype-genotype concordant patients. (A) Mutation combinations in type 3 VWD patients. The VWF:Ag, VWF:RCo and FVIII:C levels observed were below the technical error of the determination employed methods, precluding the application of statistical analyses to unravel differences between groups. (B) Mutation combinations in type 2N VWD patients. Laboratory values were found to be different (although not statistically tested) between patients with 2N mutations in both alleles (all homozygous in our cohort) versus those with a 2N mutation *in trans* with a second mutation. Patients with two 2N mutations, despite having normal VWF levels, showed a FVIII binding inability and a reduction of FVIII:C, resulting in a more severe phenotype. Null allele indicates nonsense mutations, frameshift indels and ±3 intronic mutations; HET, compound heterozygous; and HOM, homozygous.

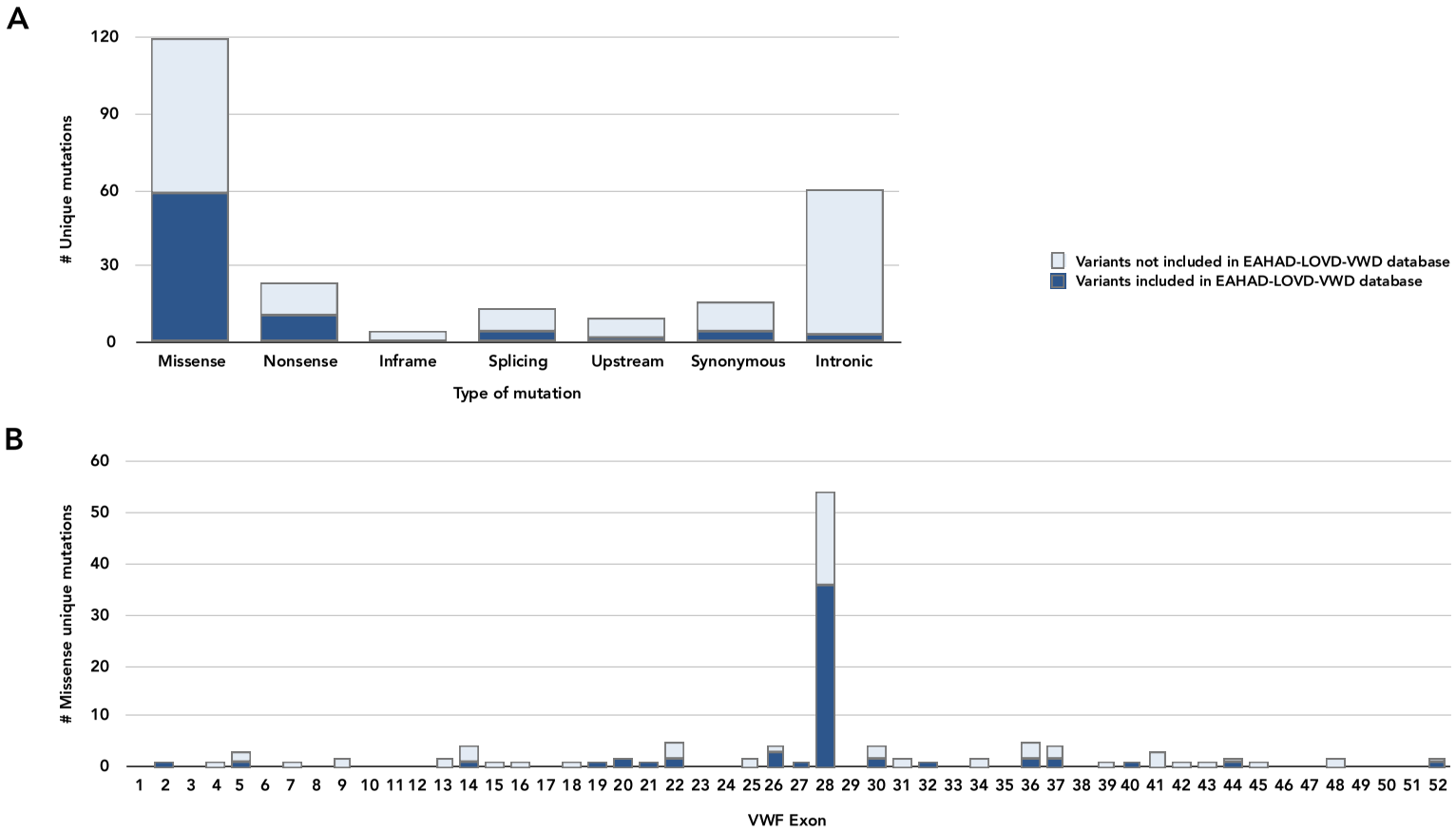


Figure S4. Graphs showing the mutation types identified in the cohort, highlighting variants that are not included in EAHAD-LOVD-VWD. (A) Number of unique mutations by type. In all cases the number of variants that are not included in EAHAD-LOVD-VWD is higher than those previously described. (B) Distribution of the 120 unique missense mutations identified along the *VWF*. Note that exon 28, the most often analyzed by Sanger sequencing, has the largest number of variants described in EAHAD-LOVD-VWD. Nonetheless, novel variants not included in EAHAD-LOVD-VWD were also found in this exon.

APPENDIX: Study group members

The members of the PCM-EVW-ES Study Group are:

Nina Borràs, Barcelona, Spain; Javier Batlle, A Coruña, Spain; Almudena Pérez-Rodríguez, A Coruña, Spain; María Fernanda López-Fernández, A Coruña, Spain; Ángela Rodríguez-Trillo, A Coruña, Spain; Esther Lourés, A Coruña, Spain; Ana Rosa Cid, Valencia, Spain; Santiago Bonanad, Valencia, Spain; Noelia Cabrera, Valencia, Spain; Andrés Moret, Valencia, Spain; Rafael Parra, Barcelona, Spain; María Eva Mingot-Castellano, Málaga, Spain; Ignacia Balda, Las Palmas de Gran Canaria, Spain; Carme Altisent, Barcelona, Spain; Rocío Pérez-Montes, Santander, Spain; Rosa María Fisac, Segovia, Spain; Gemma Iruín, Vizcaya, Spain; Sonia Herrero, Guadalajara, Spain; Sonia Herrero, Guadalajara, Spain; Inmaculada Soto, Asturias, Spain; Ángeles Rodríguez, Asturias, Spain; Nuria Fernández Mosteirín, Zaragoza, Spain; Beatriz de Rueda, Zaragoza, Spain; Víctor Jiménez-Yuste, Madrid, Spain; María Teresa Álvarez, Madrid, Spain; Nieves Alonso, Badajoz, Spain; Dolores Vilariño, Santiago de Compostela, Spain; Olga Arija, Lugo, Spain; Rosa Campos, Cádiz, Spain; María José Paloma, Navarra, Spain; Nuria Bermejo, Cáceres, Spain; Teresa Toll, Barcelona, Spain; Rubén Berrueco, Barcelona, Spain; José Mateo, Barcelona, Spain; Karmele Arribalzaga, Madrid, Spain; Pascual Marco, Alicante, Spain; Ángeles Palomo, Málaga, Spain; María Ángeles Martín, Madrid, Spain; Lizheidy Sarmiento Serrato, Madrid, Spain; María Paz Maluenda, Madrid, Spain; Belén Iñigo, Madrid, Spain; María del Mar Nieto, Jaén, Spain; Rosa Vidal, Madrid, Spain; María Paz Martínez, Ávila, Spain; Reyes Aguinaco, Tarragona, Spain; Jesús María César(†), Madrid, Spain; María Ferreiro, Pontevedra, Spain; Javier García-Frade, Valladolid, Spain; Ana María Rodríguez-Huerta, Madrid, Spain; Jorge Cuesta, Toledo, Spain; Ramón Rodríguez-González, Madrid, Spain; Faustino García-Candel, Murcia, Spain; Rosa Cornudella, Zaragoza, Spain; Carlos Aguilar, Soria, Spain; Francisco Vidal, Barcelona, Spain; and Irene Corrales, Barcelona, Spain.

(†) Jesús María César died on August 25th, 2014

SUPPLEMENTAL REFERENCES

1. Sadler JE, Budde U, Eikenboom JC, et al. Update on the pathophysiology and classification of von Willebrand disease: a report of the Subcommittee on von Willebrand Factor. J Thromb Haemost. 2006;4(10):2103-2114.
2. Montgomery RR, Christopherson P, Bellissimo DB, et al. The Complete Type I VWD Cohort Of The Zimmerman Program For The Molecular and Clinical Biology Of VWD - Phenotypic Assignment, Mutation Frequency, and Bleeding Assessment. Blood. 2013;122(21):332-332.
3. Tosetto A, Rodeghiero F, Castaman G, et al. A quantitative analysis of bleeding symptoms in type 1 von Willebrand disease: results from a multicenter European study (MCMDM-1 VWD). J Thromb Haemost. 2006;4(4):766-773.
4. Batlle J, Perez-Rodriguez A, Corrales I, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): Proposal for a new diagnostic paradigm. Thromb Haemost. 2016;114(6):
5. den Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat. 2000;15(1):7-12.
6. den Dunnen JT, Dalgleish R, Maglott DR, et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. Hum Mutat. 2016;37(6):564-569.
7. European Association for Haemophilia and Allied Disorders (EAHAD). Coagulation Factor Variant Databases. October 01, 2010 ed.
8. Berber E, James PD, Hough C, Lillicrap D. An assessment of the pathogenic significance of the R924Q von Willebrand factor substitution. J Thromb Haemost. 2009;7(10):1672-1679.
9. Rendal E, Penas N, Larrabeiti B, et al. Type 2B von Willebrand's disease due to Val1316Met mutation. Heterogeneity in the same sibship. Ann Hematol. 2001;80(6):354-360.

ARTICLE 3

Títol: "Correlació del genotip-fenotip en una cohort portuguesa de pacients incloent tot l'espectre de tipus de VWD: impacte de la NGS."

Resum

El diagnòstic de la malaltia de von Willebrand (VWD), el trastorn hemorràgic hereditari més freqüent, es caracteritza per una tendència hemorràgica variable i un fenotip heterogeni de laboratori. La seqüenciació de tota la regió codificant del *VWF* encara no s'ha convertit en una pràctica habitual als laboratoris de diagnòstic degut al seu elevat cost. No obstant això, la seqüenciació massiva (NGS) ha sorgit com una alternativa per superar aquesta limitació. Aquest treball va tenir com a objectiu determinar la correlació de genotip i fenotip en 92 individus de Portugal amb VWD de 60 famílies no relacionades. L'anàlisi genètic del *VWF* s'ha realitzat amb dues tècniques de seqüenciació directa: *Sanger* i NGS. Els resultats obtinguts amb aquestes dues aproximacions es van comparar per avaluar l'eficiència de la NGS en l'anàlisi de les mutacions en els diferents tipus de VWD. En total, es van identificar 62 mutacions diferents al *VWF*, 27 de les quals no s'havien descrit anteriorment. La NGS va detectar 26 mutacions addicionals, permetent una visió global dels al·lels mutats presents en cada tipus de VWD. Vint-i-nou casos índex (48,3%) tenien dues o més mutacions; a més, es van detectar mutacions amb efectes pleiotròpics, i la NGS va permetre una classificació adequada per a set d'ells. Tanmateix, va permetre diferenciar entre la VWD 2B i la VWD de tipus plaquetar (n = 1), la síndrome de Bernard-Soulier i la VWD 2B (n = 1), i l'hemofília A lleu i la VWD 2N (n = 2). La NGS va resultar en aproximació eficient per analitzar el *VWF* al laboratori. Aquestes troballes en la cohort de pacients portuguesos recolzen la proposta que la optimització de les estratègies de diagnòstic de la VWD, com l'anàlisi genètic, millorarà els enfocaments clínics i de laboratori, cosa que permetrà establir el tractament més adequat per a cada pacient.

Referència

Fidalgo T, Salvado R, Corrales I, Pinto SC, Borràs N, Oliveira A, Martinho P, Ferreira G, Almeida H, Oliveira C, Marques D, Gonçalves E, Diniz M, Antunes M, Tavares A, Caetano G, Kjöllerström P, Maia R, Sevivas TS, Vidal F, Ribeiro L.

Genotype-phenotype correlation in a cohort of Portuguese patients comprising the entire spectrum of VWD types: impact of NGS.

Thromb Haemost. 2016 Jul 4;116(1):17-31. doi:10.1160/TH15-07-0604. PMID: 26988807

Genotype–phenotype correlation in a cohort of Portuguese patients comprising the entire spectrum of VWD types: impact of NGS

Teresa Fidalgo[1]; Ramon Salvado[1]; Irene Corrales[2]; Silva Catarina Pinto[1]; Nina Borràs[2]; Ana Oliveira[1]; Patricia Martinho[1]; Gisela Ferreira[1]; Helena Almeida[1]; Cristina Oliveira[3]; Dalila Marques[1]; Elsa Gonçalves[1]; MJoão Diniz[3]; Margarida Antunes[3]; Alice Tavares[3]; Gonçalo Caetano[4]; Paula Kjöllerström[5]; Raquel Maia[5]; Teresa Sevivas[1]; Francisco Vidal[2]; Leticia Ribeiro[1]

[1]Centro Hospitalar e Universitário de Coimbra (CHUC) – Serviço de Hematologia Clínica, Unidade de Trombose e Hemostase, Hospital Pediátrico de Coimbra, Coimbra, Portugal; [2]Banc de Sang i Teixits (BST) – Unitat de Diagnòstic i Teràpia Molecular, Barcelona, Spain; [3]Centro Hospitalar Lisboa Norte (CHLN) – Hospital de Santa Maria – Serviço de Imunohemoterapia, Lisboa, Portugal; [4]Hospital de Faro – Serviço Hematologia, Faro, Portugal; [5]Centro Hospitalar Lisboa Central (CHLC) – Hospital D. Estefânia – Hematologia Pediátrica, Lisboa, Portugal

Summary

The diagnosis of von Willebrand disease (VWD), the most common inherited bleeding disorder, is characterised by a variable bleeding tendency and heterogeneous laboratory phenotype. The sequencing of the entire *VWF* coding region has not yet become a routine practice in diagnostic laboratories owing to its high costs. Nevertheless, next-generation sequencing (NGS) has emerged as an alternative to overcome this limitation. We aimed to determine the correlation of genotype and phenotype in 92 Portuguese individuals from 60 unrelated families with VWD; therefore, we directly sequenced *VWF*. We compared the classical Sanger sequencing approach and NGS to assess the value-added effect on the analysis of the mutation distribution in different types of VWD. Sixty-two different *VWF* mutations were identified, 27 of which had not been previously described. NGS detected 26 additional mutations, contributing to a broad overview of the mutant alleles present in each VWD type. Twenty-nine probands (48.3 %) had two or more mutations; in addition, mutations with pleiotropic effects were detected, and NGS allowed an appropriate classification for seven of them. Furthermore, the differential diagnosis between VWD 2B and platelet type VWD (n = 1), Bernard–Soulier syndrome and VWD 2B (n = 1), and mild haemophilia A and VWD 2N (n = 2) was possible. NGS provided an efficient laboratory workflow for analysing *VWF*. These findings in our cohort of Portuguese patients support the proposal that improving VWD diagnosis strategies will enhance clinical and laboratory approaches, allowing to establish the most appropriate treatment for each patient.

Keywords

Gene mutations, von Willebrand disease, von Willebrand factor, molecular biology methods, NGS

Correspondence to:
Teresa Fidalgo
Centro Hospitalar e Universitário de Coimbra (CHUC)
Serviço de Hematologia Clínica
Unidade de Trombose e Hemostase
Av Afonso Romão Coimbra 3000–602, Portugal
Tel.: +351 239 480 370, Fax: +351 239 717 216
E-mail: teresa.fidalgo@chuc.min-saude.pt

Supplementary Material to this article is available online at www.thrombosis-online.com.

Received: July 30, 2015
Accepted after major revision: March 2, 2016
Epub ahead of print: March 17, 2016

http://dx.doi.org/10.1160/TH15-07-0604
Thromb Haemost 2016; 116: 17–31

Introduction

Von Willebrand disease (VWD) is the most common inherited bleeding disorder, with a reported incidence of 0.01–1 % (1, 2). This disease is caused by a deficiency of von Willebrand factor (VWF), an adhesive glycoprotein of large dimensions with crucial functions in haemostasis derived from its ability to organise in multimers. The high-molecular-weight (HMW) multimers of VWF are essential for primary haemostasis, i. e. mediating a molecular endothelium–platelet bridge for binding collagen and the platelet receptors glycoprotein (GP)Ib and GPIIb/IIIa. In addition, VWF binds and stabilises factor VIII (FVIII) in the circulation, protecting it from rapid proteolytic degradation while delivering it to sites of vascular damage. As a result of this physiological process, VWF insufficiencies may cause haemorrhage by reducing platelet function or by reducing the FVIII concentration (1–3).

The diagnosis of VWD is based on the nature of the VWF deficiency. VWD is classified into three different types: partial or complete VWF quantitative deficiencies (types 1 and 3, respectively) and qualitative deficiency (type 2). Type 2 VWD is characterised by functional defects that lead to impaired activity and is accordingly divided into four secondary categories (2A, 2B, 2M and 2N). Functional defects lead to enhanced (2B) or reduced (2A, 2M) platelet interaction or impaired binding to FVIII (2N); structural defects due to impaired dimerisation/multimerisation can lead to type 2A/IID, 2A/IIC and 2A/IIE VWD (1, 4). Type 2A VWD can also be caused by mutations that impair the production of the larger multimers (Group I) and those with normal synthesis and

release but with increased proteolytic susceptibility to ADAMTS13 (Group II) (5). All these defects can be identified by laboratory tests that assess the functional characteristics of VWF. It is crucial to correctly diagnose the subtype of VWD because the therapeutic approaches are different for each subtype. However, this diagnosis remains difficult because the clinical and laboratory phenotypes are very heterogeneous, and a clear discrimination between some of these subtypes can be challenging (6, 7).

Molecular analysis of *VWF* is very useful in the diagnosis and classification of VWD; type 2 mutations are localised to specific functional domains in exons 18–25 (type 2N), exon 28 (types 2A, 2B and 2M) and exons 11–16, 24–26 and 51–52 (type 2A), whereas the type 1 and type 3 VWD mutations require analysis of all the essential *VWF* regions. Moreover, compound heterozygosity for *VWF* mutations in different functional domains of the VWF molecule causes pleiotropic effects and produces different phenotypes (8, 9).

To elucidate the clinical variability, which is found even in the same family, it is necessary to identify the mutational profile; however, the large size of *VWF* has been a disadvantage for studying using conventional Sanger DNA sequencing in diagnostic laboratories owing to its high costs. Nevertheless, the prospects for molecular diagnosis will be improved with the introduction of next-generation sequencing (NGS) platforms; massively parallel sequencing has reduced the cost and increased the throughput of DNA sequencing (10, 11).

With this in mind, we conducted a study with two main objectives: first, to better understand the genotype–phenotype correlation, we detailed the phenotypes of 92 Portuguese patients from 60 unrelated families with different types of VWD and *VWF* mutations; second, to decide the most appropriate strategy for VWD diagnosis in our Department of Haematology, including feasibility and costs, we compared the results obtained using conventional Sanger sequencing and NGS. This approach allowed the creation of a new algorithm that uses an efficient and cheaper methodology to establish the diagnosis, prognosis and a more accurate treatment for VWD.

Materials and methods

Patients and controls

This study included a group of 92 patients of Portuguese origin belonging to 60 apparently unrelated families; the patients were diagnosed with different types of VWD between 2007 and 2014. Sixty probands and 32 affected relatives (parents and siblings) were distributed as follows: 56 patients (31 families) regularly attended the Department of Haematology at Centro Hospitalar Universitário de Coimbra for clinical and laboratorial follow-up. The remaining 36 patients (29 families) were referred from external centres in different regions of Portugal: Lisbon (three central hospitals, one paediatric) and Faro (one central hospital). These hospitals all diagnose and treat patients with VWD and had sent samples from patients with different types of VWD to our centre for confirmation of their results, for multimer pattern analysis and for molecular study. The patients included in this study were from the central and southern regions of Portugal.

The clinical characteristics of patients were recorded through a validated questionnaire focusing on personal and family history of significant mucocutaneous haemorrhage using bleeding scores (BSs) calculated with the bleeding assessment tools (BAT) recommended by the International Society on Thrombosis and Haemostasis and Standardization Committee on von Willebrand Factor guidelines (12). In accordance with the Declaration of Helsinki, informed consent was obtained from all probands and their family members.

The 92 individuals in this study included both adults and children with a median age of 28 years (1–76) and a sex distribution of 58F : 34M. The adult probands (n = 42) had a median age of 41 years with a range of 19–76 years, and the child probands (n = 18) had a median age of 8 years with a range of 1–15 years.

Forty-two healthy volunteers acted as a control group for measurements of the PFA-100 closure time (PFA:ADP; PFA:Epi), FVIII coagulant activity (FVIII:C), VWF antigen level (VWF:Ag), ristocetin cofactor activity (VWF:RCo) and/or VWF collagen binding (VWF:CB) (▶Table 1). The control group included unrelated individuals (32F : 10M) without an individual or family history of mucocutaneous haemorrhage, with a mean age of 33.5 ± 11.12 yrs.

Samples and sample processing

Functional and immunological studies were performed on blood collected into vacuum tubes containing 3.2% sodium citrate and centrifuged within 15 minutes (min) at room temperature for 20 min at 2500 g. The obtained platelet-poor plasma was then separated into aliquots and kept frozen at –80 °C until use. The separated plasma was later subjected to coagulation and multimer assays. Genomic DNA was extracted from EDTA whole blood by automatic isolation on an iPrep™ instrument using a gDNA Blood Kit (Invitrogen, Carlsbad, USA). The DNA concentration was adjusted to a range of 25–50 ng/µl.

Algorithm for phenotypic analysis – coagulation and multimer assays

VWF analysis was performed according to an algorithm based on established recommendations (6): i) screening tests – platelet function analysis via the PFA-100 closure time (Siemens Healthcare, Marburg, Germany) in addition to measurements of the FVIII:C level, VWF:Ag level, VWF-platelet GPIb binding activity (Innovance) and VWF:RCo that were performed on a BCS XP coagulation analyser according to the manufacturer's instructions (Siemens Healthcare, Marburg, Germany); ii) specific tests – to determine disease subtype, ristocetin-induced platelet aggregation (RIPA) in platelet-rich plasma with final ristocetin concentrations of 0.5 mg/ml and 1.2 mg/ml was performed using an in-house aggregometry assay. VWF:CB was measured (except in type 3 and 2N VWD) using type III collagen (Technozym, Technoclone, Vienna, Austria), and the capacity of VWF to bind exogenous

Table 1: Demographic and phenotypic data of 60 probands at diagnosis.

	Type 1 (n = 7)	Type 3 (n = 15)	Type 2N (n = 12)	Type 2A, 2B, 2M (n = 25)	P-value
Male/Female	4/3	1/14	5/7	8/17	0.0384
Median age, years (range)	17 (1–34)	15 (1–55)	41 (13–69)	36 (2–76)	NS
Median bleeding score (range)	14 (5–18)	18 (10–23)	12.5 (2–21)	11 (1–24)	NS
PFA (COL/ADP), s PFA (COL/Epi), s	>300 >300	>300 >300	121 (99–188) 96 (82–125)	>300 >300	NS
FVIII:C* (%) (range)	26 (11–33)	2.1 (<0.25–6)	27.5 (7–38)	37.5 (15–100)	< 0.0001
VWF:Ag* (%) (range)	9.9 (3–16)	<4	68.0 (45–120)	30.0 (14–85)	< 0.0001
VWF:RCo* (%) (range)	7.2 (2–15)	<4	64.0 (47–100)	11.0 (<4–32)	< 0.0001
VWF:RCo/VWF:Ag* (range)	0.8 (0.67–1.00)	-	1.0 (0.83–1.11)	0.4 (0.14–0.73)	< 0.0001
VWF:CB* (%) (range)	4.5 (2–16)	-	-	13.0 (<4–53)	< 0.0001

Normal range: PFA (COL/ADP), 93–161 s; PFA (COL/Epi), 67–99 s; FVIII:C, 49–149 %; VWF:Ag, 50–160 %; VWF:RCo, 48–173 % and VWF:CB, 60–130 %. PFA-platelet function analysis; s – seconds; *Median and range; Chi-squared testing for categorical variables; One-way ANOVA followed by Bartlett's test for all linear variables; NS, no significance.

FVIII (VWF:FVIIIB) (Stago, Asnieres, France) was evaluated using ELISA. The multimer pattern was evaluated by SDS-agarose gel electrophoresis (low-resolution gel electrophoresis using 0.9 % low-gelling temperature agarose or medium resolution using 1.6 % agarose), followed by Western blotting and detection with rabbit anti-human VWF antibody (Dako, Glostrup, Denmark) using the alkaline phosphatase staining method (13).

The potential probands with VWF deficiency were diagnosed by evaluating the following criteria: presence of reduced (< 30 %) VWF:RCo levels; a discrepant VWF:RCo/VWF:Ag ratio of < 0.7 that suggested type 2 VWD; a discrepant FVIII:C/VWF:Ag ratio of < 0.5 that suggested type 2N VWD; the absence of VWF (or presence of only trace amounts) and a FVIII:C below 5 % that distinguished type 3 VWD from severe type 1 VWD (1).

The rate of *VWF* mutations found in type 1 VWD decreases when the VWF level is > 30 % (1). To avoid high costs, in our approach, patient samples with borderline VWF levels were not pursued for molecular analysis and Sanger sequencing.

Alloantibodies to VWF

To test alloantibodies to VWF in type 3 patients, an in-house ELISA-based assay was used as previously described (14).

Strategy for *VWF* mutation analysis

The molecular analysis in this study used two different direct sequencing methodologies:

i) Sanger direct sequencing: For the seven years of this study, the following sequences were determined using an approach based on VWD subtypes: type 1 severe and type 3 VWD – all coding *VWF*; types 2B and 2M VWD - only exon 28; type 2A – first exon 28 and then exons 11–16, 24–26 and 51–52 if there was no mutation in exon 28; and type 2N – exons 18–25. To confirm differential diagnosis between VWD 2B and platelet type VWD (PT-VWD), exon 2 of *GP1BA*.

ii) NGS: The coding *VWF* [regions of interest (ROIs)] were analysed in all samples (DNA samples were sent to Unitat de Diagnòstic i Teràpia Molecular, Banc de Sang i Teixits (BST, Barcelona, Spain)).

Sequencing of *VWF* using NGS and identification of genetic variants

A recently described automated method for the Access Array™ platform (Fluidigm, South San Francisco, CA, USA) derived from a previous NGS protocol (10, 15) was used for *VWF* amplification. Briefly, the resulting products from all patients included in this study were pooled and simultaneously sequenced in a MiSeq Desktop Sequencer (Illumina, San Diego, CA, USA) run. Barcoded sequences were de-multiplexed and individually analysed. The ROIs included in our NGS design were approximately 1300 bp of the *VWF* promoter region, exons 1–52 and the intronic flanking regions (at least 20 bp). The depth of coverage and primer design is described in detail in the Suppl. Material (available online at www.thrombosis-online.com). The NGS pipeline output, in paired sequence files (FASTQ format), was used as an input for the analysis, which began with CLC Genomic Workbench software (Qiagen, Hilden, Germany) and then proceeded to VariantStudio (Illumina). The optimal analysis parameters (coverage, minor allele counts, percent of variant alleles, etc.) were adjusted to obtain the optimal performance for variant detection (see Suppl. Material, available online at www.thrombosis-online.com). This workflow allows the alignment of the resulting sequences against the human genome sequence (hg19) and concurrent *in silico* analysis, permitting the identification of potential pathogenic variants, discriminating pseudogene sequences and filtering the known

polymorphisms described to date in the dbSNP (16) and 1000 Genomes databases (17).

For confirmation of the variants identified by NGS, the corresponding region was amplified using PCR and sequenced by Sanger direct sequencing, as described in the Suppl. Material (available online at www.thrombosis-online.com).

Multiplex ligation-dependent probe amplification (MLPA)

DNA samples from patients lacking identified pathogenic variants using PCR/direct sequencing, were screened using SALSA MLPA P011 and P012 VWF kits (version B2; MRC-Holland) to detect deletions/duplications in *VWF*. Fragment size analysis was performed using an ABI 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Data normalisation was performed using four healthy controls and Coffalyser. Net software (MRC-Holland) was used for analysis using the fragment analysis files (.fsa) obtained from the Applied Biosystems Genetic Analyzer as input.

Mapping of the *VWF* exon 31 deletion breakpoints

The exact locations of the deletion breakpoints in one type 1 VWD proband in whom the mutation was initially identified by MLPA were determined by primer walking followed by Sanger sequencing. The genomic region from the end of exon 30 to intron 31 was amplified using five primer sets. The following primer pair sequences were used in the specific assay for the detection of the deletion: forward primer 5'-GCTGCAGTATGGAAGCATCA-3' and reverse primer 5'-GAGTTGAGGTGAGGCTGGAG-3'.

In silico analysis

The impact of missense changes was further analysed according to the recommended criteria that included evolutionary conservation of an amino acid or nucleotide, location and context within the protein sequence and the biochemical consequence of the amino acid substitutions (18, 19). This analysis was assessed using five different *in silico* algorithms: PROVEAN, SIFT, PolyPhen-2, MutationAssessor and MutationTaster (Suppl. References, available online at www.thrombosis-online.com) (last accessed September 21, 2015). Missense variants were considered deleterious if at least three of the five prediction programs suggested a pathogenic effect.

Potential disease-causing variants due to aberrant splicing were evaluated using a set of five different *in silico* algorithms: Human Splicing Finder (HSF), MaxEntScan, NNSPLICE, NetGene2 and Splice View (Suppl. References, available online at www.thrombosis-online.com) (last accessed September 21, 2015). This approach improves the accuracy of the splicing analysis of intronic, deep intronic and exonic variants that may directly disrupt constitutive splice sites or indirectly create cryptic splice sites (20, 21). MutPred Splice was used to confirm probable exonic splicing variants (22). A local splice site effect was considered potentially deleterious when at least three (intronic variants) or four (exonic variants) of these predictions for the variant vicinity were significant.

Genetic databases

As recommended by the guidelines (18, 21, 23), we defined those variants that are not reported in the international databases (population- and disease-specific) as well as in published literature as 'novel'. Therefore, in addition to using the polymorphism databases (dbSNP and 1000 Genomes databases) mentioned above, we checked variants for inclusion in the Exome Aggregation Consortium (ExAC) (24), Exome Variant Server (EVS) (25), Human Gene Mutation Database (26) and VWF international mutation databases (27–29).

Assessment of the pathogenicity of variants

The following criteria were used to evaluate the pathogenicity of the variants: 1) whether the variant was a stop/frameshift variant, which was considered to most likely be disease causing, 2) co-segregation in the family, 3) whether the variation had been previously identified in international databases, 4) *in silico* evaluation and 5) presence of the second mutant allele in the case of autosomal recessive inheritance.

Variants were classified as pathogenic, likely pathogenic, uncertain significance, likely benign or benign based on the available evidence, according to the practice guidelines for the evaluation of pathogenicity recently published by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (18).

The amino acid numbering and nomenclature used is according to the international recommendations for the description of sequence variants of the Human Genome Variation Society (http://www.HGVS.org) and applied to genetic variants in haemostasis (30).

Statistical analysis

Differences between groups were analysed by Chi-squared testing for categorical variables and one-way ANOVA followed by Bartlett's test for all linear variables. Tests were assumed significant whenever the two-tailed p-value was < 0.05. These statistical analyses were performed using GraphPad Prism 6.0 for Windows (GraphPad Software, La Jolla, CA, USA; www.graphpad.com).

Results

Laboratory haemostasis findings

The present study included a cohort of 60 probands from different regions of Portugal with severe/intermediate types of VWD. According to the study algorithm, in the first approach, a standardised BS and VWF:Ag, VWF:RCo and FVIII:C levels were taken into account to diagnose these patients. BS showed an overall inverse correlation with the plasma levels of VWF:Ag ($p = 0.0022$, Spearman's rho = 0.3822), VWF:RCo ($p = 0.0013$, Spearman's rho = 0.3991) and FVIII:C ($p = 0.0020$, Spearman's rho = 0.3829).

Forty-eight probands presented with low levels of VWF:RCo (9.15 ± 9.19 %), and 12 presented with a reduced FVIII:C/VWF:Ag

Table 2: Phenotypic and molecular data of seven probands with severe type 1 VWD.

Laboratory values							Molecular data				
Probands ID (n)	Gender/ age (M/F year)	Bleeding Score	FVIII:C (%)	VWF: Ag (%)	FVIII:C/VWF: Ag ratio	VWF: RCo (%)	VWF: CB (%)	Exon/ Intron	Nucleotide Change	Amino Acid Change	Domain	Genotype
P17 (1)	M/24	18	11	3	3.7	2	2	43	**c.7400A>C**	**p.Gln2467Pro**	C3	Homozygous
P19 (1)	M/27	12	26	16	1.6	15	16	31	**c.5312–104_5455+642del**		A3	Heterozygous
P65 (1)	F/34	6	26	10	2.6	7	6	38	c.6699_6702dupAGGC	p.Cys2235Argfs*8	D4	Heterozygous
P18 (1)	M/5	4	33	11	3.0	9	NA	5	**c.440A>G**	**p.Gln147Arg**	D1	Compound heterozygous
								42	c.7086C>A ‡	p.Cys2362*	C2	
P15 (1)	F/5	16	31	7	4.4	5	3	15	c.1892C>T	p.Ala631Val	D2	Heterozygous
								46/45	c.7730–4C>G‡		C4	Homozygous
P12 (1)	M/10	16	27	6	4.5	6	3	38	c.6699_6702dupAGGC	p.Cys2235Argfs*8	D4	Compound heterozygous
								43	c.7437G>A‡	p.(=)	C3	
P48 (1)	F/5	5	17	13	1.3	10	12	29	**c.5140G>C**	**p.Ala1714Pro**	A3	Compound heterozygous

NA- not available; the mutations marked in **bold** were not previously reported. ‡ indicates mutation in *trans*.

ratio (0.36 ± 0.15); these features prompted a detailed phenotypic study and mutational analysis.

The phenotypic characteristics of the probands were mainly distributed as follows: severe type 1 VWD (n = 7), type 3 VWD (n = 15) and type 2 VWD (n = 38). Specific tests indicated the subdivision of type 2 VWD as follows: type 2A (n = 9), type 2B (n = 6), type 2M (n = 11) and type 2N (n = 12) (▶Table 1). Familial studies (n = 32) included 25 affected relatives, three type 3 carriers and four type 2N carriers (data not shown).

The NGS method confirmed all variants detected by the Sanger method (100% sensitivity) and detected 26 additional variants: six missense, three synonymous and 17 intronic [15 of 17 (88%) were deep intronic]. In total, 62 different variants, whose frequency in the different populations studied in 1000 Genomes was below 1%, are summarised in Suppl. Table 1 (available online at www.thrombosis-online.com): 30 missense variants (48.5%), three nonsense (4.8%), two small deletions (3.2%), one large deletion (1.6%), one duplication (1.6%), five synonymous (8%) and 20 intronic variants (32.3%). The frequencies of those variants were also checked in the ExAC and EVS_EA population databases, which revealed differing values for some variants (Suppl. Table 1, available online at www.thrombosis-online.com).

In total, 27 variants had never been reported in the population databases and international *VWF* databases (Suppl. Table 1, available online at www.thrombosis-online.com).

Prediction of pathogenic variants

Twenty-six out of 31 missense variants were predicted to be deleterious (Suppl. Table 2, available online at www.thrombosis-online.com). Fifteen of 56 variants (missense, synonymous and intronic) were detected using some in silico tools with probable impact on splicing; however, only six reached the required score (≥ 3 for intronic and ≥ 4 for exonic mutations) and were considered potential splice site variants (Suppl. Table 3, available online at www.thrombosis-online.com). According to the practice guidelines (18), these *in silico* analyses combined with other evidence data (population, functional and segregation) permitted the following classification of the variants: pathogenic (n = 29; 46.9%), likely pathogenic (n = 4; 6.4%), uncertain significance (n = 4; 6.4%), likely benign (n = 18; 29%) and benign (n = 7; 11.3%) (Suppl. Table 1, available online at www.thrombosis-online.com). Accordingly, for the classification of these variants, we used the term 'mutation' only to refer to pathogenic variants.

Variants were identified in all probands of this cohort (100%), and 29 (48.3%) had two or more variants along *VWF*.

Potential functional impact of novel mutations

Eleven of the 27 novel variants (41%) were pathogenic: five missense changes (c.440A>G, p.Gln147Arg; c.2637C>A,

p.Asp879Glu; c.4117G>T, p.Asp1373Tyr; c.5140G>C, p.Ala1714Pro; and c.7400A>C, p.Gln2467Pro); two nonsense mutations (c.4666C>T, p.Gln1556*; c.7086C>A, p.Cys2362*); two small deletions (c.100delT, p.Arg34Aspfs*49; c.5414_5415del, p.Val1805Glyfs*8); one large deletion [c.5312–104-_5455+642del (890 bp)] and one potential splice site mutation at the consensus 5'-GT donor splice site (c.1533+1G>A) (Suppl. Table 1, available online at www.thrombosis-online.com). The remaining 16 variants accounted for two missense changes that were signalled as benign variants (c.3590C>A, p.Pro1197Gln; c.3686T>C; p.Val1229Ala) and 14 deep intronic variants that were considered with uncertain significance (2) and likely benign (12) (Suppl. Table 1, available online at www.thrombosis-online.com).

Phenotype–genotype correlations

The probands' phenotypic and molecular data were separated according to the VWD type, and only the variants with predicted pathogenicity (mutations) are summarised in ▶Tables 2–6. For four probands, the phenotype-based classification was changed after mutation analysis.

Type 1 VWD

Seven probands had a severe type 1 VWD profile with very low VWF:Ag and VWF:RCo levels but with moderately reduced FVIII:C levels (▶Table 2). In fact, five of these seven patients had an increased FVIII:C/VWF:Ag ratio with a median value of 3.7 (range 2.6–4.5). Thirteen variants were distributed along several regions of *VWF* with alterations in domains D1, D2, A3, D4, C2, C3 and C4 (▶Figure 1). As shown in ▶Table 2, 10 mutations (five not previously described) were associated with this phenotype. All probands except three had compound heterozygous mutations: one had a novel homozygous missense mutation, p.Gln2467Pro, located in the C3 domain, another had a novel heterozygous large deletion of exon 31 located in the A3 domain, and the third had a

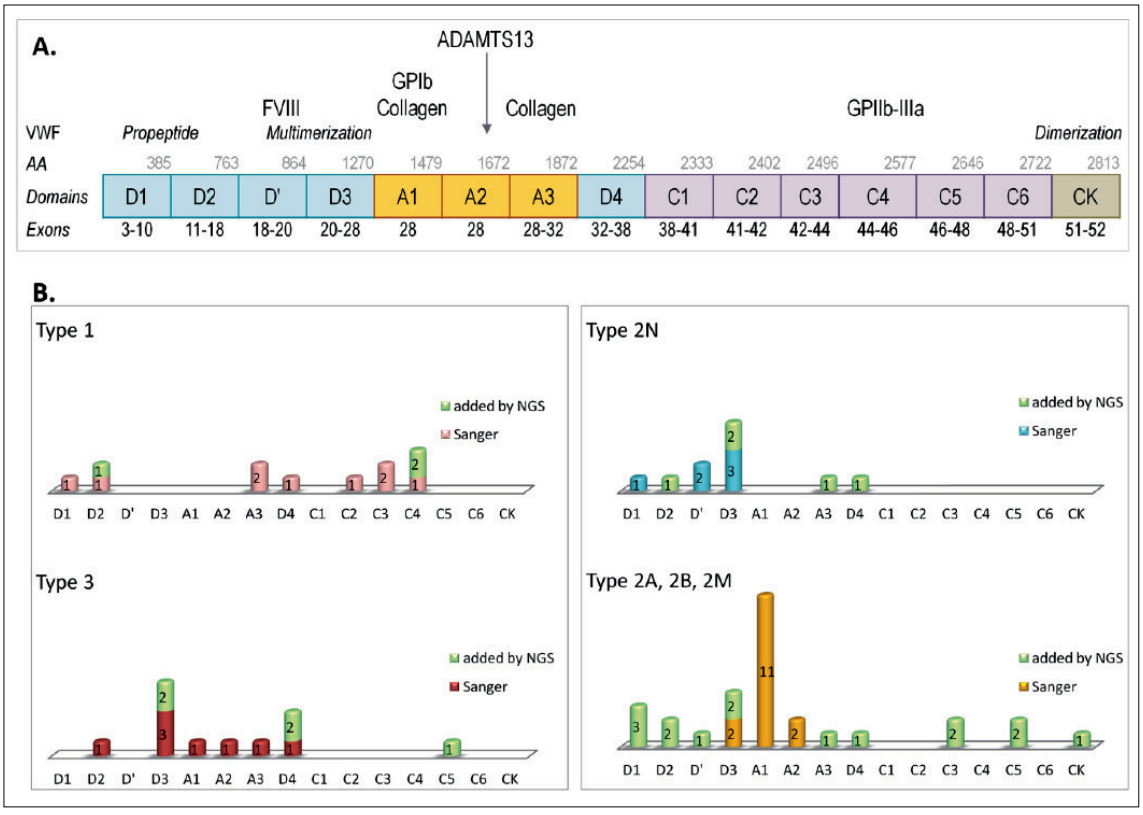


Figure 1: Distribution of 62 different *VWF* mutations identified in 60 probands with VWD. A) The scheme represents *VWF* and the corresponding domains of the VWF protein (55, 56). B) The graphics show the mutation distribution associated with the *VWF* domains for severe type 1, type 3, type 2N and type 2A, 2B, 2M VWD. Each graphic discriminates between the mutations previously identified by Sanger sequencing and the mutations added by NGS. No evident cluster of alterations was observed for type 1 and type 3 VWD, and the spreading along *VWF* was accentuated. For type 2N VWD, a mutation cluster in the region coding for D'-D3 was preserved. For type 2A, 2B and 2M VWD, mutation spreading was evident, but a mutation cluster emerged in the region coding for the A1 domain.

duplication mutation c.6699_6702dupAGGC, p.Cys2235Argfs*8. MLPA analysis detected the large deletion, and PCR followed by direct DNA sequencing mapped the intronic breakpoints c.5312–104_5455+642del, indicating that the deletion was 890 bp in size (Suppl. Figure 1, available online at www.thrombosis-online.com).

The remaining four probands showed compound heterozygosity in *trans* for missense and null mutations (nonsense and splice site). Two probands had a recessively inherited missense mutation p.[Gln147Arg];[Cys2362*] and [p.Ala631Val];[c.7730–4C>G] located in the VWF propeptide (VWFpp) region (D1–D2 domain). Both had a typical multimeric pattern with a heavy predominance of VWF dimers (IIC). The proband P12 had a duplication mutation c.6699_6702dupAGGC, p.Cys2235Argfs*8 and the synonymous mutation c.7437G>A in the splice region, resulting in severely decreased VWF:Ag levels. The same duplication mutation was detected in the homozygous state in type 3 VWD patients (►Table 3). Finally, the proband P48 had a compound heterozygosity in *trans* for the novel missense mutation p.Ala1714Pro in the A3 domain and the potential splice site variant c.5170+10C>T. This missense mutation in the proband's mother was inherited as a dominant trait with a type 2M VWD phenotype, normal multimeric pattern and low VWF:CB/VWF:Ag ratio (0.54), thus showing a pleiotropic effect.

Type 3 VWD

The severity of 15 probands with type 3 VWD was evidenced by the low median age (15 years) and higher BSs, which were consistent with the low levels of FVIII:C (median 2%, range < 0.25–6), VWF:Ag (< 4%) and VWF:RCo (< 4%) and the absent VWF multimeric pattern (►Table 1 and ►Table 3). Mutations were identified in 14 probands, 13 of whom were homozygous for only one mutation (93%). In total, 13 different variants and six mutations were scattered throughout VWF (►Figure 1).

Three novel mutations were identified in our cohort of type 3 VWD patients: the nonsense mutation c.4666C>T (p.Gln1556*), the small deletion c.5414_5415delTG (p.Val1805Glyfs*8) and the splice site mutation c.1533+1G>A.

The majority of the probands (13/15; 87%) were homozygous for a null mutation that comprised the nonsense mutations c.3931C>T (p.Gln1311*) and c.4666C>T (p.Gln1556*), the small deletion c.5414_5415delTG (p.Val1805Glyfs*8) and the duplication c.6699_6702dupAGGC (p.Cys2235Argfs*8). One proband was compound heterozygous for [p.Tyr1146Cys];[c.1533+1G>A]; a study of the parents allowed the tracing of the mutation inheritance in each allele: the father was heterozygous for p.Tyr1146Cys, showing a mild type 1 VWD, whereas the mother was heterozygous for c.1533+1G>A but was asymptomatic.

The remaining type 3 VWD proband and her brother were homozygous for a dinucleotide change c.3485_3486delinsTG that results in the variant p.Pro1162Leu, which affects the D3 domain; however, the real contribution of this variant to the severe quantitative phenotype must still be clarified. Thus, causative mutations for type 3 VWD were found in 14 out of 15 probands (93%).

Table 3: Phenotypic and molecular data of 14 probands with type 3 VWD.

Laboratory values						Molecular data			
Probands ID (n)	**Gender/age (M/F year)**	**Bleeding Score**	**FVIII:C* (%)**	**VWF: Ag* (%)**	**VWF:RCo* (%)**	**Exon/ Intron**	**Nucleotide Change**	**Amino Acid Change**	**Domain**	**Genotype**
P2, P3, P4, P5, P63, (5)	F/21; M/24; F/15; F/1; F/21	21, 17,11, NA, 10	<2.5 (<2.5–4)	<4	<4	28	c.3931C>T	p.Gln1311*	A1	Homozygous
P6, P7, P8, P9 (4)	F/13;F/12; F/41; F/46	21, 19, 22,13	<2.5 (<2.5–2.6)	<4	<4	28	**c.4666C>T**	**p.Gln1556***	A2	Homozygous
P10 (1)	F/55	17	6	3	<4	31	**c.5414_5415delTG**	**p.Val1805Glyfs*8**	A3	Homozygous
P11, P13, P14 (3)	F/2; F/14; F/8	14, 22, 23	2.2 (2–6)	<4	<4	38	c.6699_6702dupAGGC	p.Cys2235Argfs*8	D4	Homozygous
P64 (1)	F/1	NA	<4	<4	<4	26	c.3437A>G	p.Tyr1146Cys	D3	Heterozygous
						13/13	c.1533+1G>A‡		D2	Heterozygous

*Median and range; NA- not available; the mutations marked in **bold** were not previously reported. ‡ indicates mutation in *trans*.

Laboratory values						Molecular data				
Probands ID (n)	**Gender/age (M/F year)**	**Bleeding Score**	**FVIII:C* (%)**	**VWF:Ag* (%)**	**VWF:RCo * (%)**	**VWF: FVIIIB (%)**	**Exon/ Intron**	**Nucleotide Change**	**Amino Acid Change**	**Domain**	**Genotype**
P54 (1)	F/50	7	7	100	100	NA	19	c.2446C>T	p.Arg816Trp	D'	Homozygous
P52 (1)	F/51	21	11	52	47	0.49	19	c.2446C>T	p.Arg816Trp	D'	Heterozygous
							3	**c.100delT‡**	**p.Arg34Aspfs*49**	D1	Heterozygous
P51; P55, P56, P58, P60, P61 P53 (7)	M/41; F/31;F/61; F/23; F/69; M/40 F/43	3; 10;14; 8;21;16 15	27.5 (19–33)	74.5 (45–88)	70.5 (50–93)	0–2	20	c. 2561G>A	p.Arg854Gln	D'	Homozygous
P57 (1)	F/35	2	18	46	50	NA	20	c. 2561G>A	p.Arg854Gln	D'	Homozygous
							39	c.6890C>T‡	Pro2297Leu	C1	Heterozygous
P59 (1)	M/40	11	38	66	68	6	20	c.2561G>A	p.Arg854Gln	D'	Heterozygous
							20	**c.2637C>A‡**	**p.Asp879Glu**	D3	Heterozygous
P62 (1)	M/13	6	32	120	100	NA	19	c.2451T>A	p.His817Gln	D'	Heterozygous
							26	c.3485_3486delinsTG ‡	p.Pro1162Leu	D3	Homozygous

* Median and range; NA- not available; the mutations marked in **bold** were not previously reported. ‡ indicates mutation in *trans*.

Table 4: Phenotypic and molecular data of 12 probands with type 2N VWD.

Among those patients who were homozygous for the c.3931C>T (p.Gln1311*) mutation, one proband (P2) showed an ineffective response after replacement therapy, and the presence of an inhibitory alloantibody against VWF was confirmed by an anti-VWF ELISA assay. On the basis of these data, the patient received FVIII bypass therapy according to their clinical situation.

Type 2N VWD

All the 12 type 2N VWD probands had reduced FVIII:C levels with median levels of 27.5% (range 7–38), normal or near normal median VWF:Ag levels of approximately 68% (range 45–120), normal VWF:RCo levels of 64% (range 47–100) and a normal VWF multimeric pattern (▶Table 1). The reduced FVIII:C/VWF:Ag ratio (0.4; 0.07–0.6) was suggestive of type 2N VWD, which was confirmed by reduced VWF:FVIII binding and/or molecular studies.

In total, 11 different variants and seven mutations were found in the *VWF* regions affecting the domains D1, D', D3, A3 and D4. Despite this distribution, five missense mutations were clustered in exons 19 and 26 of *VWF,* which encode the FVIII-binding D'–D3 domains (▶Figure 1). The frequently described missense mutations c.2561G>A (p.Arg854Gln) and c.2446C>T (p.Arg816Trp) repetitively occurred in our cohort, being the most recurrent mutations (75% of cases) (▶Table 4).

The missense mutation p.Arg816Trp was found in the two probands who had the lowest FVIII:C levels (▶Table 4), one in a homozygous state and the other compound heterozygous with a null allele p.[Arg816Trp];[Arg34Aspfs*49] (2N/3). Eight out of nine individuals were homozygous for p.Arg854Gln, and one was compound heterozygous for a novel mutation, c.2637C>A (p.Asp879Glu).The FVIII:C/VWF:Ag ratio was not significantly different in homozygotes or heterozygotes for the p.Arg854Gln mutation. In a 13-year-old African male (P62) who was previously diagnosed with mild haemophilia A (based only on functional studies), heterozygosity for the missense variant c.2451T>A (p.His817Arg) and homozygosity for the variant c.3485_3486delinsTG (p.Pro1162Leu) were detected in the D' and D3 domains, respectively. The two variants, p.Pro1162Leu and p.His817Arg, located in the VWF region, which affects the FVIII binding capacity, could both contribute to the type 2N phenotype.

In this group of 2N VWD patients, two probands (P61 and P62) had previously been misclassified as having mild haemophilia A.

Type 2B VWD

The six probands classified as 2B VWD had a very low median VWF:RCo level of 11% (range 10–24) and a VWF:RCo/VWF:Ag

ratio of 0.4 (0.35–0.50); the VWF multimeric patterns showed partial or complete loss of the HMW multimers (►Table 5). Among the six probands, four had thrombocytopenia (15, 50, 79 and 130 × 10^3 platelets/μl). Because the samples from three probands were provided by other centres, it was only possible to perform RIPA in the other three patients who showed an enhanced platelet aggregation in response to 0.5 mg/ml of ristocetin.

The frequently described mutations associated with VWD 2B, p.Arg1306Trp and p.Val1316Met, in the region coding for the A1 VWF domain, were found in five patients (►Table 5 and ►Figure 1). One proband (P25) was referred to us with a suspected diagnosis of Bernard–Soulier syndrome because of the severe thrombocytopenia (15 × 10^3 platelets/μl) and the giant platelets observed in a peripheral blood smear. However, the coagulation multimer assays and molecular studies allowed a reclassification as type 2B VWD with a compound heterozygou*s* in *cis* for mutations p.[Arg1306Trp;Arg854Gln]. The proband P50 was compound heterozygous in *trans* for p.Val1316Met with two other variants, the novel benign missense variant p.Pro1197Gln and c.5170+10C>T.

Finally, in proband P66, despite the VWF levels being consistent with type 2B VWD, no mutation was found in exon 28 *VWF*. The phenotype data included low VWF:RCo levels (47%), a VWF:RCo/VWF:Ag ratio of 0.4, partial loss of the HMW multimers and RIPA at 0.25 mg/ml. A differential diagnosis of probable platelet-type VWD (PT-VWD) was performed using simple RIPA mixing studies (control platelets in patient plasma), which revealed the absence of aggregation with 0.5 mg/ml of ristocetin. Therefore, a platelet-type phenocopy (pseudo-VWD) was considered. The sequencing of *GP1BA* revealed the presence of a previously described missense mutation, c.751G>T, p.Asp251Tyr in the heterozygous state in exon 2, confirming the reclassification of this patient as PT-VWD (Suppl. Figure 2, available online at www.thrombosis-online.com).

Type 2A and 2M VWD

The VWF:RCo and VWF:Ag levels corresponded to discrepant VWF:RCo/VWF:Ag ratios (0.39 ± 0.22) in these two groups of patients. The VWF:CB levels were similar to the VWF:RCo levels for type 2A VWD and to the VWF:Ag levels for 2M VWD. The VWF multimeric patterns in the type 2A VWD group showed the loss of the HMW multimers, and the pattern for the 2M VWD group showed the presence of all–size multimers (►Table 6). Probands with type 2A VWD had a significantly higher BS (median = 15) than those with type 2M VWD (median = 8.5) (p = 0.020, Mann–Whitney U test).

In 20 probands, 10 different mutations were detected in exon 28, which encodes the VWF A1 and A2 domains (►Figure 1). Among these five mutations (p.Arg1374Cys p.Arg1374His, p.Arg1399Cys, p.Ser1506Leu and p.Ile1628Thr) that occurred repetitively, p.Ile1628Thr was the most frequent mutation in 2A VWD (40% of cases) and p.Arg1374His was the most frequent in 2M VWD (37% of cases) (►Table 6).

Of the nine probands with a 2A VWD phenotype, only one (P20) was heterozygous for the missense mutation in exon 26 (domain D3), c.3388T>C p.Cys1130Arg, with low levels of VWF:RCo (16%), a discrepant VWF:RCo/VWF:Ag ratio of < 0.7 and a characteristic 2A (IIE) VWF multimeric pattern (absence of large multimers and no triplet structure of individual bands), which is indicative of a multimerisation defect (►Table 6, Suppl. Figure 3, available online at www.thrombosis-online.com).

The probands P23 and P24 were heterozygous for two mutations affecting the same position and located in the first amino acid of the A1 loop domain: c.3814T>G, p.Cys1272Gly and c.3815G>T, p.Cys1272Phe. Interestingly, the mutation p.Cys1272Phe showed compound heterozygosity with a potential splice site variant (c.7288–68G>A) that could strongly activate a cryptic donor (Suppl. Tables 1 and 3, available online at www.thrombosis-online.com). Despite this mutation location being less common in type 2A VWD, the laboratory phenotype was consistent in these two probands: a marked decrease in VWF:RCo levels (5%–7%), a VWF:RCo/VWF:Ag ratio of < 0.18, a multimeric pattern showing a loss of high- and intermediate-molecular weight multimers (Suppl. Figure 3, available online at www.thrombosis-online.com) and RIPA at 1.25 mg/ml was absent.

The remaining six probands were heterozygous for two frequently described missense mutations clustered in the A2 domain: c.4517C>T, p.Ser1506Leu (n = 2) and c.4883T>C, p.Ile1628Thr (n = 4). The two patients heterozygous for the p.Ser1506Leu mutation, described as a VWD 2A group I defect with intracellular proteolysis of large VWF multimers, showed a similar multimeric pattern (Suppl. Figure 3, available online at www.thrombosis-online.com). These probands had moderate/severe type 2A VWD with low levels of VWF:Ag (median 17.5%) and VWF:RCo (median 11%), a low VWF:RCo/VWF:Ag ratio (median 0.64) and no RIPA at 1.25 mg/ml. The remaining four probands were heterozygous for p.Ile1628Thr, which has been described as a VWD 2A group II defect characterised by hypersensitivity to ADAMTS13 with increased proteolysis in plasma (5). They had mild/moderate type 2A VWD with low levels of VWF:Ag (median 54.8%) and VWF:RCo (median 13.2%), a low VWF:RCo/VWF:Ag ratio (median 0.2) and normal RIPA at 1.25 mg/ml. These four patients showed a typical proteolytic pattern with a lack of large VWF multimers and the presence of a pronounced triplet structure, which is characteristic of group 2A/II.

Of the 11 probands with the type 2M VWD phenotype, eight (72%) were heterozygous for the frequent missense mutations involving two arginine residues in the A1 domain: p.Arg1315Cys, p.Arg1315His, p.Arg1374Cys and p.Arg1374His (►Table 6).

The laboratory phenotype of the probands heterozygous for p.Arg1315Cys and p.Arg1315His mutations was similar with low levels of VWF:Ag (median 19%) and VWF:RCo (median 11.5%), a decreased VWF:RCo/VWF:Ag ratio (0.66) and a normal multimeric pattern. However, the proband with the mutation p.Arg1315His was compound heterozygous with the silent mutation c.7464C>T detected in exon 44, which may create a new donor splicing site (►Table 6 and Suppl. Table 3, available online at www.thrombosis-online.com). The proband's mother, who was also heterozygous for p.Arg1315His, had a type 1 VWD

Table 5: Phenotypic and molecular data of five probands with type 2B VWD and one with PT-VWD.

Laboratory values										
Probands ID (n)	**Gender/ age (M/F year)**	**Bleeding Score**	**FVIII:C* (%)**	**VWF: Ag* (%)**	**VWF: RCo* (%)**	**VWF: CB (%)**	**VWF:RCo/ VWF:Ag Ratio**	**Platelet X10³/µl**	**RIPA 0.5 mg/ml**	**VWF Multimers**
P26, P27, P28 (3)	F/41; F/11; F/30	20,8, 5	47 (23–78)	25 (20–30)	11 (10–12)	12 (9–16)	0.44 (0.40–0.50)	79, 230, 242	NA, E NA	LHMWM
P25 (1)	F/41	5	26	29	11	13	0.38	15	E	LHMWM
P50 (1)	F/11	10	37	60	24	22	0.35	130	NA	LHMWM
P66 (1)	M/40	1	101	122	47	50	0.4	50	E	Partial LHMWM

*Median and range; RIPA – ristocetin-induced platelet aggregation; NA – not available; E – enhanced; LHMWM – loss of high-molecular-weight multimers; □ indicates mutation in *cis*.

phenotype with higher levels of VWF:Ag (32%) and VWF:RCo (26%).

The missense mutations p.Arg1374Cys and p.Arg1374His were the most frequent cause of type 2M VWD in our cohort of patients. All the six probands showed a decreased VWF:RCo/VWF:Ag ratio (median 0.3), a high VWF:CB/VWF:Ag ratio and the presence of all–size multimers, albeit with a certain degree of smearing. Two probands had the missense mutation c.4195C>T; p.Arg1399Cys associated with a mild phenotype and pronounced smearing of the multimer pattern.

The last patient (P38) was compound heterozygous for four previously undescribed missense variants. However, a family study was not conducted; therefore, which of them were in *cis* or in *trans* was unknown: p.[Val343Leu(;)Ile482Met(;)Asp1373Tyr(;)-His1419Gln]. Two variants (p.Asp1373Tyr and p.His1419Gln) were clustered in the A1 domain, and the other two (p.Val343Leu and p.Ile482Met) were in the VWFpp region (D1–D2 domain). Of these four variants, only one was pathogenic (p.Asp1373Tyr), whereas the other three were benign variants. The phenotype was mild with a normal multimeric pattern (▶Table 6).

Table 6: Phenotypic and molecular data of 20 probands with type 2A and 2M VWD.

Laboratory values									
Probands ID(n)	**Gender/ age (M/F year)**	**Bleeding Score**	**FVIII* (%)**	**VWF: Ag (%)***	**VWF: RCo(%)***	**VWF:RCo/ VWF:Ag Ratio**	**VWF: CB (%)**	**RIPA 1.25 mg/ml**	**Multimers/ subtype**
P20 (1)	F/36	19	27	22	16	0.7	10	Decresead	LHMWM/ 2A/IIE
P23 (1)	F/58	15	22	31	7	0.23	8	Absent	LHMWM/ 2A
P24 (1)	M/22	10	23	22	3	0.14	4	Absent	LHMWM/ 2A
P42, P43 (2)	M/46; F/8	14;24	28; 37	21;14	12; 10	0.57; 0.71	4;10	Decresead	LHMWM/ 2A/I
P44, P45, P46, P47 (4)	F/24; M/65 F/46; F/28	12; 15 17;12	40 (34–49)	52 (42–65)	13 (10–23)	0.22 (0.15–0.35)	11 (6–19)	Normal	LHMWM/ 2A/II
P30 (1)	F/36	7	48	16	11	0.69	19	Decresead	Normal/ 2M
P29 (1)	F/6	3	24	22	14	0.64	NA	NA	Smear/ 2M
P32, P33 (2)	F/48; F76	9;1	45; 87	70; 36	20; 13	0.17; 0.36	NA, 32	NA	Smear/ 2M
P34, P35 P36, P37 (4)	M/2; M/73 F/63;F/31	4; 9;3	37 (15–60)	26 (20–57)	8 (4–10)	0.3 (0.16–0.50)	20 (16–45)	NA	Normal/ 2M
P39, P40 (2)	F/42; M/9	12;7	100;82	85; 55	31; 32	0.36; 0.58	53; 37	Normal	Smear/ 2M
P38 (1)	M/32	7	51	31	19	0.61	40	Normal	Normal/ 2M

*Median and range; RIPA – ristocetin-induced platelet aggregation; NA – not available; LHMWM – loss of high-molecular-weight multimers; the mutation marked in **bold** was not previously reported; ‡ indicates mutation in *trans*.

Molecular data			
Exon/ Intron	**Nucleotide Change**	**Amino Acid Change**	**Domain**
28	c.3916C>T	p.Arg1306Trp	A1
28	c.3916C>T	p.Arg1306Trp	A1
20	c.2561G>A	p.Arg854Gln	D'
28	c.3946G>A	p.Val1316Met	A1
2, *GP1BA*	c.751G>T	p.Asp251Tyr (Asp235Tyr)	

The knowledge of all putative mutations for every patient of this cohort permitted a broad overview of the inheritance and the combination of mutant alleles that affected each VWD type. The recessive inheritance of a majority of the mutant alleles identified in types 1, 3 and 2N VWD matched the following combination: missense/null in 5/7 for severe type 1; null/null in 13/15 for type 3 and missense/missense in 11/12 for type 2N. As expected, type 2 VWD (2A, 2B and 2M) showed a dominantly inherited trait with the allelic combination wild-type/missense in 23/25 for type 2 (►Tables 2–6; ►Figure 2).

Molecular data			
Exon/ Intron	**Nucleotide change**	**Amino Acid Change**	**Domain**
26	c.3388T>C	p.Cys1130Arg	D3
28	c.3814T>G	p.Cys1272Gly	A1
28	c.3815G>T	p.Cys1272Phe	A1
28	c.4517C>T	p.Ser1506Leu	A2
28	c.4883T>C	p.Ile1628Thr	A2
28	c.3944G>A	p.Arg1315His	A1
44	c.7464C>T‡	p.(=)	C3
28	c.3943C>T	p.Arg1315Cys	A1
28	c.4120C>T	p.Arg1374Cys	A1
28	c.4121G>A	p.Arg1374His	A1
28	c.4195C>T	p.Arg1399Cys	A1
28	**c.4117G>T**	**p.Asp1373Tyr**	A1

Discussion

This study comprised a genotype–phenotype correlation in a cohort of 60 unrelated Portuguese families (92 individuals) diagnosed with types 1, 2 and 3 VWD. The VWD diagnosis based on the results of a battery of laboratory tests is often difficult, and therefore, establishing patients' disease severity and risk of bleeding becomes challenging (6, 7). Thus, in our diagnosis laboratory, a phenotype and genotype analysis in patients with VWD has become crucial for understanding the VWF function and great inter-individual variability.

As shown in ►Figure 1, besides the additional mutations identified by NGS, the characteristic mutation distribution across *VWF* was preserved in each VWD type: in type 3 VWD, the mutations were scattered throughout *VWF*; in severe type 1 VWD, the profile was similar, but the absence of mutations in the regions coding for the A1, A2, D' and D3 domains was evident; in type 2N VWD, a mutation cluster in the region coding for the D1, D' and D3 domains was detected; in types 2A, 2B and 2M VWD, the mutation spreading was notable, but a mutation cluster in the regions coding for the A1 domain emerged. These findings are in line with previous reports (5, 31–34).

The group of seven probands with severe type 1 VWD showed a high degree of genetic heterogeneity with 10 different mutations, of which five (three missense mutations, one nonsense and one large deletion) have not been previously described (35%). The *in silico* analysis for all three missense mutations showed that they were expected to be deleterious and disease causing (Suppl. Table 1, available online at www.thrombosis-online.com). Five out of seven probands with mutations in regions coding for the VWF D1, D2, D4 and C2-C4 domains showed an increased FVIII:C/VWF:Ag ratio, which reflected impaired VWF secretion, possibly due to misfolding, as previously described (35).

The missense mutation p.Ala631Val has only been described in a large study of healthy controls (36) as an outlier owing to a mild type 1 VWD phenotype being presented. Nevertheless, in the present study, this mutation was found in compound heterozygosity with c.7730–4C>G, and the inheritance of both mutations was responsible for a severe phenotype (Suppl. Figure 4, available online at www.thrombosis-online.com). In addition, the compound heterozygosity of the novel missense mutation p.Ala1714Pro with c.5170+10C>T was associated with severe type 1 VWD. Moreover, when inherited alone, this mutation exhibited a type 2M VWD phenotype (Suppl. Figure 4, available online at www.thrombosis-online.com). The p.Ala1714Pro mutation as well as others in the A3 domain showed an accentuated defective collagen binding with a normal multimer pattern (37).

The variants c.7730–4C>G and c.5170+10C>T, although the low probability of a deleterious effect predicted by *in silico* algorithms, they seem to have had some influence on VWD phenotype in these families: the first one showed effect in autosomal recessive pattern and the second one showed a cumulative effect in a compound heterozygous state on decrease of VWF levels (Suppl. Figure 4, available online at www.thrombosis-online.com). However, the variant c.5170+10C>T is frequent in the studied populations,

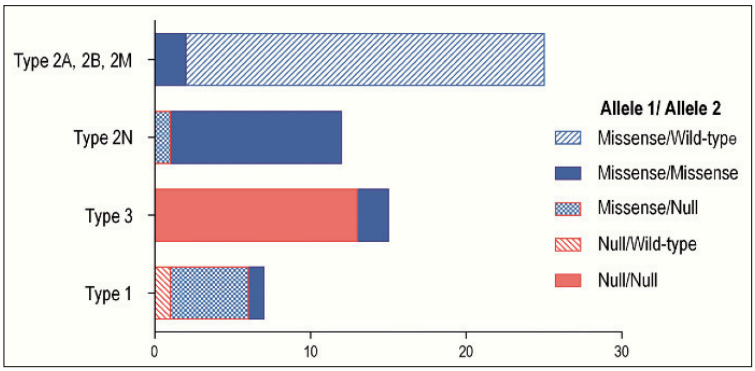


Figure 2: Mutation combination in mutated alleles in different VWD types. The main type of mutation detected in each VWD type was as follows: severe type 1 VWD, missense/null; type 3 VWD, null/null; 2N VWD, missense/missense; and type 2 VWD (2A, 2B and 2M), missense/wild type.

despite the variable frequencies found in different databases (ExAC Aggregated_Populations: MAF = 0.0036; EVS_EA: MAF = 0.0046 and dbSNP: MAF = 0.0089) and it may be controversial to consider this a pathogenic variant. Nevertheless, it should be noted that other deep intronic changes that can only be detected by whole gene sequencing could be present and may have also influenced these phenotypes. These findings corroborate the concept that the variants' pathogenicity should be determined by the entire body of evidence in each family (18). Moreover, these likely pathogenic variants could not explain the disease; however, they seem to influence the phenotype and could behave as a pathogenic in one case and non-pathogenic in another, depending on the other co-inherited changes (15, 18). These data have proven the added value of practice guidelines for the evaluation of the pathogenicity of potential splice site variants. Among the six variants that were predicted as potential splice sites by *in silico* analysis, two were considered pathogenic, two likely pathogenic and the other two with uncertain significance (Suppl. Table 1, available online at www.thrombosis-online.com).

Therefore, these classification criteria did not replace the functional mRNA studies, the only method that can elucidate the true effect of the variation detected (38); however, this classification provided a more precise genotype–phenotype correlation.

Our cohort of type 3 VWD evidenced, as usual, two patterns of mutations: homozygous for null mutations and compound heterozygous/homozygous for missense mutations of VWF (►Figure 2). Two nonsense mutations (p.Gln1311*, p.Gln1556*) and a novel small deletion (p.Val1805Glyfs*8) were expected to have the obvious deleterious mechanisms, i.e. premature termination codons and a frameshift leading to a truncated VWF protein. The homozygous p.Gln1311* mutation associated with a gene conversion and initially reported in Spanish Romani families (39) was the most frequent mutation in 5/15 families, of which one was also Romani. Among these five patients, one had developed an inhibitor that has also been described in homozygous gene conversions (40). In three families studied in this cohort, obligate carriers of p.Gln1556* and p.Gln1311* were not phenotypically silent and manifested mild VWF levels and mucocutaneous bleeding symptoms, showing an inheritance pattern that is co-dominant (33) and not recessive, as described in other studies. One carrier of p.Gln1556* had the variant p.Tyr1584Cys in *trans* without any worsening of phenotype. The small duplication p.Cys2235Argfs*8 previously described in Spanish patients (10) and also repetitively found in this Portuguese study could share a common ancestral origin.

The molecular mechanisms of missense mutations in compound heterozygous state in the D3 domain (p.Tyr1146Cys) could be related to ineffective multimerisation and the subsequent defective intracellular transport and secretion with intracellular retention (41).

Finally, the p.Pro1162Leu in the homozygous state detected in a type 3 VWD proband was also found in another proband with type 2N VWD (►Table 4). This controversial finding implied that p.Pro1162Leu was unlikely to be a pathogenic variant for type 3 VWD. However, the same genotype was also the only one found in a type 3 VWD Spanish family (15); therefore, it seems reasonable to expect that the p.Pro1162Leu (c.3485_3486delinsTG) variant is a part of a compound genotype that causes type 3 VWD, which has not yet been entirely explained. This possibility is in accordance with the lack of detection of molecular mechanism of type 3 VWD despite NGS improvements. The explanation for this includes deep intronic mutations that may only be identified through whole-gene sequencing and distant regulatory elements outside VWF (33).

The functional VWF:FVIIIB defect (type 2N) was mainly caused by the missense mutation p.Arg854Gln in the homozygous state in our cohort, similarly to previous studies. This is a frequent deleterious variant in Caucasian populations (9, 42). Nevertheless, genetic variability was introduced by compound heterozygous with novel mutations: p.Arg854Gln with the deleterious novel missense mutation p.Asp879Glu and p.Arg816Trp with the novel small deletion c.100delT (p.Arg34Aspfs*49) (►Table 4, Suppl. Table 1, available online at www.thrombosis-online.com). The ethnic variability was noted in an African proband who showed heterozygosity for the frequent polymorphism p.His817Gln in association with p.Pro1162Leu in the homozygous state. Recent multiethnic studies found that VWF missense variants in the D' and D3 domains, previously identified in European ancestry VWD probands, are more frequent in African Americans (AAs) (36, 43, 44).

These variants include p.His817Gln, which was strongly associated with diminished FVIII:C levels (43). The p.Pro1162Leu variant (caused by c.3485_3486delinsTG) was previously only reported in one healthy AA in the heterozygous state (36). Nevertheless, the same variant caused by a different nucleotide change (c.3485C>T) was described with polymorphic frequency in AAs (MAF = 0.262) (16, 17). Taken together these data, it could be reasonable to interpret that such different variants at the genomic level, although theoretically lead to identical amino acid change, could result in unlike consequences at the transcriptional level and to distinct deleterious effects.

This pleiotropic effect on the VWF molecule may be explained by the ethnic diversity of the VWF coding polymorphisms recently documented in large cohorts (36, 43, 44). As previously reported, these mutations that result in low VWF antigen levels could be partially compensated by the high baseline VWF antigen levels found in Africans. They could be VWD causing only when they are confined in ethnic-specific haplotypes; in fact, only a study of large Caucasian and African cohorts could link specific VWF variants to a bleeding phenotype (44).

In agreement with previous studies, the genotype–phenotype correlation in the 25 patients with types 2A, 2B and 2M VWD was explained by the VWF domain affected and was therefore divided into structural and functional defects (41, 45).

Our type 2B VWD cohort of patients with gain-of-function mutations in the A1 domain showed two common mutations, p.Arg1306Trp and p.Val1316Met (46). In addition, we observed a highly variable platelet count with a large range (15–242 × 103 platelets/µl), which is consistent with previous studies (►Table 5), although only four patients had thrombocytopenia. This variability was explained by several mutations and their coding positions within the VWF A1 domain, e.g. the altered VWF GPIb-α–binding conformation (34). Although p.Arg1306Trp and p.Val1316Met are frequently correlated with thrombocytopenia, we observed some heterogeneity even in individuals with the same mutation. These data were also in agreement with the wide degree of heterogeneity of the clinical and laboratory features reported for affected members of families that have type 2B VWD (34).

Herein, we report a case of PT-VWD misdiagnosed as type 2B VWD. The functional studies allowed a differential diagnosis, but molecular studies confirmed the recently described missense mutation p.Asp251Tyr (Asp235Tyr) in *GP1BA* (47). PT-VWD is certainly an underdiagnosed deficiency, where molecular study is a valuable diagnostic tool.

We found mutations related to all previously described type 2A structural defects (5, 31, 48, 49): i) a multimerisation defect (p.Cys1130Arg in domain D3) with the absence of large multimers and no triplet structure; ii) a change in protein folding (p.Cys1272Gly/Phe in loop of A1 domain) that causes of loss of large and intermediate multimers; iii) intracellular proteolysis of large VWF multimers (p.Ser1506Leu) with a loss of large multimers; and iv) hypersensitivity to ADAMTS13 (p.Ile1628Thr in A2 domain) with the absence of large VWF multimers and increased triplet structure (►Table 6, Suppl. Figure 3, available online at www.thrombosis-online.com).

The mutations p.Cys1130Arg and p.Ser1506Leu showed disproportionate VWF:RCo/VWF:Ag ratios, but near 0.7; the mutations p.Cys1272Gly/Phe and p.Ile1628Thr had very low ratio values, indicative of enhanced proteolysis, as previously described (5, 48) (Suppl. Figure 3, available online at www.thrombosis-online.com).

Our findings in type 2M VWD patients support the well-characterised profile: mutations in the A1 domain, which are typically associated with decreased or absent RIPA and a low VWF:RCo/VWF:Ag ratio, combined with a normal VWF:CB/VWF:Ag ratio. We identified the usual A1 domain mutation clusters, p.Arg1315Cys/His and p.Arg1374Cys/His, with the addition of the novel mutations p.Asp1373Tyr and p.His1419Gln. The VWD subtype classification of the p.Arg1374His mutation has been controversial. This mutation was described as type 2A (49) or type 2M (50). In fact, some authors argue that the mutation is difficult to classify as type 2M (because of the possibly relative decrease in large VWF multimers) or any other type 2 (because of the normal banding pattern of each multimer) VWD; therefore, a classification of type 2U VWD (unclassifiable) has been suggested (51). In our study, three probands showed a multimer pattern, with the full complement of multimers, and one showed a slight decrease in the largest forms. Given the criterion of a low VWF:RCo/VWF:Ag ratio and normal VWF:CB/VWF:Ag ratio, which was evident in each of the four probands, they were classified as having type 2M VWD.

The same criterion was applied to the p.Arg1315Cys, p.Arg1374Cys and p.Arg1399Cys mutations that showed a smeary multimer pattern (no clear separation between individual oligomer triplets), which was particularly pronounced in two probands and two relatives with p.Arg1399Cys. This structural VWF alteration usually occurs when the responsible mutation involves cysteine residues, which may affect the disulphide bonding of VWF dimers (9, 49). A recent study analysed the impact of cysteine mutations that affect the carboxyl-terminal domains of VWF (52), it clearly demonstrated the importance of cysteine residues for the structural conformation and consequently multimerisation, even in patients with type 2A VWD with a smeary multimer pattern. Therefore, the smeary appearance could be associated with the alterations in VWF domains that determine whether VWD is type 2A or 2M; however, the cysteine mutations seem to be the hallmark of this multimer pattern (9, 49, 52).

This study described, for the first time, the mutational spectrum in a cohort of Portuguese VWD patients. Overall, VWD studies highlight the high degree of variability in clinical presentation and the considerable heterogeneity of the molecular basis, which supports the great interest shown in reports from different geographic regions (8, 33, 50, 53, 54). Moreover, large studies have highlighted the ethnic variability in the phenotype of many VWF missense variants (36, 43, 44). We applied these insights to the context of familial studies to determine their usefulness for predicting individual bleeding risk. Accordingly, the genotype–phenotype correlation in each proband family with diverse VWD types was assessed to establish family overviews and elucidate phenotypic discrepancies.

This accurate analysis was only possible because NGS has made it easier to study VWF ROIs. The NGS methodology applied in this study was the same as that used in the recently published 'Molecular and clinical profile of von Willebrand disease in Spain (PCM–EVW–ES)' project. NGS has proven to be an excellent technology that enables more rapid diagnosis with a huge economic advantage (only 70€/sample, which is even cheaper than most phenotypic tests) (15). Similarly, these findings support the adjustment of our VWD diagnosis algorithm, introducing the complete sequencing of VWF (NGS) when the VWF:RCo level is < 30 % or the FVIII/VWF:Ag ratio is < 0.5.

Because of the geographical proximity of the two populations, we expected to detect the same common mutations in the patients of our cohort as for the cohort of PCM–EVW–ES (15). In fact, in type 3 VWD patients, the following were observed: the nonsense mutation p.Gln1311* was the most prevalent, and the missense mutation p.Pro1162Leu was also present in homozygous state. The p.Cys2235Argfs*8 was only detected in Spanish and Portuguese patients, and interestingly, the variant p.His817Gln was also noted in association with diminished FVIII:C levels and the type 2N phenotype.

Our new molecular study approach permitted the identification of 27 novel VWF mutations, with some occurring repetitively, illustrating the advantages of identifying the most prevalent mutations in a region and their correlation with the severity of bleeding phenotypes. The approach allowed us to distinguish between clinical situations that have the same symptoms with different genetic causes such as mild Haemophilia A and type 2N VWD, Bernard–Soulier syndrome and 2B VWD as well as 2B VWD and PT-VWD. In contrast, despite being a restricted study, we identified some pleiotropic mutation effects. In both situations, it was evinced that molecular studies are indispensable for an accurate diagnosis.

The molecular characterisation of VWD patients allows precise classification into the correct VWD type and the identification of carriers in familial genetic studies. This classification is particularly relevant and even mandatory for genetic counselling for type 3 VWD and, in general, for patients with higher bleeding risk. Moreover, it facilitates the evaluation of prophylactic requirements and clinical orientation, particularly in risky situations.

In conclusion, this study of 60 VWD Portuguese families will contribute to the better understanding of the molecular genetics of *VWF*-related phenotypes. NGS, in our experience, provides an effective laboratory workflow for the analysis of a single large gene, such as *VWF*. These findings support the idea that improving VWD diagnosis strategies will enhance clinical and laboratory approaches; thus, the most appropriate treatment for each patient can be determined.

Acknowledgements

We are grateful to the Forum Hematológico for the funding for this project. Development of *VWF* analysis by NGS was supported in part by a Baxter BioScience Grant (H13–000845). This work was also supported by the Spanish Ministerio de Economía y Competitividad (MINECO)-Instituto de Salud Carlos III (ISCIII) (PI1201494 and RD12/0042/0053). We like also thank to all the medical and patients who contributed to this study.

Conflicts of interest

None declared.

What is known about this topic?

- VWD is an inherited bleeding disorder characterised by a variable bleeding tendency and heterogeneous laboratory phenotype.
- A genotype–phenotype correlation in patients with VWD has become crucial for understanding the function and great inter-individual variability of VWF.
- With the advent of NGS, this correlation is easier to obtain because of the possibility of sequencing the entire *VWF* coding region for a very reasonable cost.

What does this paper add?

- Twenty-seven novel *VWF* mutations or potential mutations, including seven missense mutations, two nonsense mutations, two small deletions and a large deletion of exon 31, extend the mutational spectrum of *VWF*.
- The genotype–phenotype correlation analysis in our cohort of patients enables the unravelling of several diagnostic discrepancies and the identification of potential pleiotropic effects of mutations.
- The present study showed that NGS provides an accurate molecular analysis; however, its interplay with a detailed clinical data registry and familial studies is crucial.

References

1. Sadler JE, Budde U, Eikenboom JCJ, et al. Update on the pathophysiology and classification of von Willebrand disease: a report of the Subcommittee on von Willebrand Factor. J Thromb Haemost 2006; 4: 2103–2114.
2. Batlle J, Pérez-Rodríguez A, López-Fernández MF. Classification of von Willebrand disease. In: Federici AB et al, editor. Von Willebrand Dis Basic Clin Asp 1st ed. Oxford: Wiley-Blackwell; 2011. pp. 74–85.
3. Sadler JE. von Willebrand factor assembly and secretion. J Thromb Haemost 2009; 7 (Suppl 1): 24–27.
4. Schneppenheim R. The pathophysiology of von Willebrand disease: therapeutic implications. Thromb Res 2011; 128 (Suppl): S3–7.
5. Michiels JJ, van Vliet HHDM. Dominant von Willebrand disease type 2A groups I and II due to missense mutations in the A2 domain of the von Willebrand factor gene: diagnosis and management. Acta Haematol 2009; 121: 154–166.
6. Castaman G, Hillarp A, Goodeve A. Laboratory aspects of von Willebrand disease: test repertoire and options for activity assays and genetic analysis. Haemophilia 2014; 20 (Suppl 4): 65–70.
7. Budde U, Favaloro E A. Laboratory diagnosis of von Willebrand disease: the phenotype. In: Federici AB et al, editor. Von Willebrand Dis Basic Clin Asp 1st ed. Oxford: Wiley-Blackwell; 2011. pp. 100–113.
8. Corrales I, Ramírez L, Altisent C, et al. Rapid molecular diagnosis of von Willebrand disease by direct sequencing. Detection of 12 novel putative mutations in VWF gene. Thromb Haemost 2009; 101: 570–576.
9. Goodeve AC. The genetic basis of von Willebrand disease. Blood Rev 2010; 24: 123–134.
10. Corrales I, Catarino S, Ayats J, et al. High-throughput molecular diagnosis of von Willebrand disease by next generation sequencing methods. Haematologica 2012; 97: 1003–1007.

11. Goodeve AC. Current avenues in Von Willebrand disease research. Haematologica 2014; 8: 47–52.
12. Rodeghiero F, Castaman G, Tosetto A. How I treat von Willebrand disease. Blood 2009; 114: 1158–1165.
13. Pérez-Rodríguez A, García-Rivero A, Lourés E, et al. Autosomal dominant C1149R von Willebrand disease: Phenotypic findings and their implications. Haematologica 2009; 94: 679–686.
14. Siaka C, Rugeri L, Caron C, et al. A new ELISA assay for diagnosis of acquired von Willebrand syndrome. Haemophilia 2003; 9: 303–308.
15. Batlle J, Pérez-Rodríguez A, Corrales I, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM – EVW – ES): Proposal for a new diagnostic paradigm. Thromb Haemost 2015; 2: 1–11.
16. National Center for Biotechnology Information dbSNP (build 137). Available from: http://www.ncbi.nlm.nih.gov/snp. Accessed July 14, 2015.
17. 1000 Genomes. A deep catalog of human genetic variation. Available from: www.1000genomes.org. Accessed July 14, 2015.
18. Richards S, Aziz N, Bale S, et al. ACMG Standards and Guidelines Standards and guidelines for the interpretation of sequence variants : a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015; 17: 405–424.
19. Gonzalez-Garay ML. The road from next-generation sequencing to personalized medicine. Per Med 2014; 11: 523–544.
20. Jian X, Boerwinkle E, Liu X. In silico tools for splicing defect prediction: a survey from the viewpoint of end users. Genet Med 2014; 16: 497–503.
21. Wallis Y, Payne S, Mcanulty C, et al. Practice Guidelines for the Evaluation of Pathogenicity and the Reporting of Sequence Variants in Clinical Molecular Genetics. Available from: http://www.acgs.uk.com/media/774853/evaluation. Accessed September 21, 2015.
22. Mort M, Sterne-Weiler T, Li B, et al. MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing. Genome Biol 2014; 15: R19.
23. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. Nature 2014; 508: 469–476.
24. Exome Aggregation Consortium (ExAC), Cambridge, MA. Available from: http://exac.broadinstitute.org/. Accessed September 21, 2015.
25. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA. Available from: http://evs.gs.washington.edu/EVS/. Accessed September 21, 2015.
26. HGMD Human gene mutation database (HGMD® Professional) from BIOBASE Corporation. Available from: www.biobase-international.com/hgmd. Accessed September 21, 2015.
27. International Society on Thrombosis and Haemostasis-Scientific and Standardisation Committee of von Willebrand Factor (ISTH-SSC). VWF Online Database (VWFdb). Available from: http://www.vwf.group.shef.ac.uk/vwd.html. Accessed September 21, 2015.
28. European Association for Haemophilia and Allied Disorders (EAHAD). Coagulation Factor Variant Databases. Available from: https://www.grenada.lumc.nl/LOVD2/VWF/home.php?select_db=VWF. Accessed September 21, 2015.
29. Hemobase. Available from: http://vwf.hemobase.com/. Accessed September 21, 2015.
30. Goodeve A, Reitsma P, Mcvey J. Nomenclature of genetic variants in hemostasis. J Thromb Haemost 2011; 9: 852–855.
31. Gadisseur A, Berneman Z, Schroyens W, et al. Laboratory Diagnosis of von Willebrand Disease Type 1/2E (2A Subtype IIE), Type 1 Vicenza and Mild Type 1 Caused by Mutations in the D3, D4, B1–B3 and C1–C2 Domains of the von Willebrand Factor Gene. Acta Haematol 2009; 121: 128–138.
32. Michiels JJ, Gadisseur A, van der Planken M, et al. Laboratory and molecular characteristics of recessive von Willebrand disease type 2C (2A subtype IIC) of variable severity due to homozygous or double heterozygous mutations in the D1 and D2 domains. Acta Haematol 2009; 121: 111–118.
33. Bowman M, Tuttle A, Notley C, et al. The genetics of Canadian type 3 von Willebrand disease: Further evidence for co-dominant inheritance of mutant alleles. J Thromb Haemost 2013; 11: 512–520.
34. Federici AB, Mannucci PM, Castaman G, et al. Clinical and molecular predictors of thrombocytopenia and risk of bleeding in patients with von Willebrand disease type 2B : a cohort study of 67 patients. Blood 2011; 113: 526–534.
35. Eikenboom J, Hilbert L, Ribba AS et al. Expression of 14 von Willebrand factor mutations identified in patients with type 1 von Willebrand disease from the MCMDM-1VWD study. J Thromb Haemost 2009; 7: 1304–1312.
36. Bellissimo DB, Christopherson PA, Flood VH, et al. VWF mutations and new sequence variations identified in healthy controls are more frequent in the African-American population. Blood 2012; 119: 1–3.
37. Riddell AF, Gomez K, Millar CM, et al. Characterization of W1745C and S1783A: 2 novel mutations causing defective collagen binding in the A3 domain of von Willebrand factor. Blood 2009; 114: 3489–3496.
38. Corrales I, Ramírez L, Altisent C, et al. The study of the effect of splicing mutations in von Willebrand factor using RNA isolated from patients' platelets and leukocytes. J Thromb Haemost 2011; 9: 679–688.
39. Casaña P, Martínez F, Haya S, et al. Q1311X: a novel nonsense mutation of putative ancient origin in the von Willebrand factor gene. Br J Haematol 2000; 111: 552–555.
40. Surdhar GK. Homozygous gene conversion in von Willebrand factor gene as a cause of type 3 von Willebrand disease and predisposition to inhibitor development. Blood 2001; 98: 248–250.
41. Schneppenheim R, Budde U. von Willebrand factor: the complex molecular genetics of a multidomain and multifunctional protein. J Thromb Haemost 2011; 9 (Suppl 1): 209–215.
42. Jacquemin M. Factor VIII-von Willebrand factor binding defects in autosomal recessive von Willebrand disease type Normandy and in mild hemophilia A. New insights into factor VIII-von Willebrand factor interactions. Acta Haematol 2009; 121: 102–105.
43. Johnsen JM, Auer PL, Morrison AC, et al. Common and rare von Willebrand factor (VWF) coding variants, VWF levels, and factor VIII levels in African Americans: the NHLBI Exome Sequencing Project. Blood 2013; 122: 590–597.
44. Wang QY, Song J, Gibbs R, et al. Characterizing polymorphisms and allelic diversity of von Willebrand factor gene in the 1000 Genomes. J Thromb Haemost 2013; 11: 261–269.
45. Meyer D, Fressinaud E, Mazurier C. Clinical, laboratory, and molecular markers of type 2 von Willebrand disease. In: Federici AB, Lee CA, Berntorp EE, Lillicrap D, Montgomery RR E, editor. Von Willebrand Dis Basic Clin Asp Oxford: Wiley-Blackwell; 2011. pp. 137–147.
46. Gadisseur A, Hermans C, Berneman Z, et al. Laboratory diagnosis and molecular classification of von Willebrand disease. Acta Haematol 2009; 121: 71–84.
47. Enayat S, Ravanbod S, Rassoulzadegan M, et al. A novel D235Y mutation in the GP1BA gene enhances platelet interaction with von Willebrand factor in an Iranian family with platelet-type von Willebrand disease. Thromb Haemost 2012; 108: 946–954.
48. Woods I, Sanchez-Luceros A, Kempfer A, et al. C1272F: A novel type 2A von Willebrand's disease mutation in A1 domain; its clinical significance. Haemophilia 2012; 18: 112–116.
49. Budde U, Schneppenheim R, Eikenboom J, et al. Detailed von Willebrand factor multimer analysis in patients with von Willebrand disease in the European study, molecular and clinical markers for the diagnosis and management of type 1 von Willebrand disease (MCMDM-1VWD). J Thromb Haemost 2008; 6: 762–771.
50. Castaman G, Federici A, Tosetto A, et al. Different bleeding risk in type 2A and 2M von Willebrand disease: A 2-year prospective study in 107 patients. J Thromb Haemost 2012; 10: 632–638.
51. Gadisseur A, van der Planken M, Schroyens W, et al. Dominant von Willebrand disease type 2M and 2U are variable expressions of one distinct disease entity caused by loss-of-function mutations in the A1 domain of the von Willebrand factor gene. Acta Haematol 2009; 121: 145–153.
52. Yadegari H, Driesen J, Pavlova A, et al. Insights into pathological mechanisms of missense mutations in C-terminal domains of von Willebrand factor causing qualitative or quantitative von Willebrand disease. Haematologica 2013; 98: 1315–1323.
53. Hampshire DJ, Abuzenadah AM, Cartwright A, et al. Identification and characterisation of mutations associated with von Willebrand disease in a Turkish patient cohort. Thromb Haemost 2013; 110: 264–274.
54. Yadegari H, Driesen J, Pavlova A, et al. Mutation distribution in the von Willebrand factor gene related to the different von Willebrand disease (VWD) types in a cohort of VWD patients. Thromb Haemost 2012; 108: 662–671.
55. Zhou Y-F, Eng ET, Zhu J, et al. Sequence and structure relationships within von Willebrand factor. Thombosis Hemost 2012; 120: 449–458.
56. Lenting PJ, Christophe OD. von Willebrand factor biosynthesis, secretion, and clearance: connecting the far ends. Blood 2015; 125: 2019–2029.

Material Suplementari Article 3

Supplementary Material to Fidalgo et al. "Genotype–phenotype correlation in a cohort of Portuguese patients comprising the entire spectrum of VWD types: impact of NGS" (Thromb Haemost 2016; 116.1)

Suppl. Materials and Methods

Sanger direct sequencing

Forty-seven primer pairs were used to amplify the 52 exons and the flanking splice regions of *VWF*. The primer sequences and polymerase chain reaction (PCR) conditions were previously described (8). The following primer pair was used to study exon 2 of *GP1BA:* 5' CTGGAGAATCTCGACACCCTTC; 5' AGTGTGAAATTTGGGGTCCATCT. After amplification, the PCR products for all samples were sequenced using an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and Big Dye Terminator v1.1 Cycle Sequencing Chemistry (Applied Biosystems), according to the protocols recommended by the manufacturer. Data from the Sequence Analysis Software was aligned against the native *VWF* sequence (GenBank no. NC_000012.10) using the SeqScape® v2.5 software (Applied Biosystems) for mismatch detection.

NGS VWF sequencing - depth of coverage and primers design

The minimum depth of coverage required for each nucleotide in the ROI to be identified/considered a candidate mutation was >=10X. However, the mean coverage for all the ROIs (although highly variable) was at least 500-fold coverage for each base in all patients. All designed primers were evaluated to corroborate the absence of SNPs in the primer binding sequences (particularly in the 3' region) that could result in preferential or single allele amplification. In addition, the primers were designed to be highly specific to avoid pseudogene (*VWFP1*) amplification. However, the optimal Fluidigm PCR conditions were not stringent enough, and approximately 4–5 % of the reads obtained for each patient corresponded to pseudogene sequences. The specific alignment and segregation of reads corresponding to *VWF* and *VWFP1* was achieved by adjusting the parameters of the sequence analysis software (15).

Suppl. Table 1. Sixty- two different variants were identified in 60 families (92 individuals) with VWD. Distribution by the type of variant with assessment of the pathogenicity using different types of data: population, *in silico*, function and segregation.

Nucleotide Change	Amino Acid Change	Exon	MAF[a] (ExAC)	MAF[a] (EVS_EA)	rs ID	HGMD	VWF databases	*In silico* score[b] (Supp.Table 2)	VWD Subtype associated	Familial Segregation	Variants classification	Familial Studies Prop./relatives	Sanger	NGS
Missense variants														
c.440A>G	**p.Gln147Arg**[d]	5	ND	ND	NA	ND	ND	3	Type 1[d]	No	Pathogenic	1/0	·	·
c.1027G>T	p.Val343Leu	9	0.000008294	ND	NA	ND	ND	4	Co-inh./2M	No	Benign	1/0		·
c.1446C>G	p.Ile482Met	13	0.0001272	ND	rs569669757	ND	ND	4	Co-inh./2M	No	Benign	1/0		·
c.1892C>T	p.Ala631Val[d]	15	0.0002834	0.00029	rs199963222	ND	ND	3	Type 1[d]	Yes	Likely Pathogenic	1/2	·	·
c.2446C>T	p.Arg816Trp	19	0.00002489	ND	rs121964894	DM	Yes	5	Type 2N	No	Pathogenic	2/0	·	·
c.2561G>A	p.Arg854Gln	20	0.003081	0.005581	rs41276738	DM	Yes	3	Type 2N	Yes	Pathogenic	11/5	·	·
c.2637C>A	**p.Asp879Glu**	20	ND	ND	NA	ND	ND	5	Type 2N	Yes	Pathogenic	1/2	·	·
c.3388T>C	p.Cys1130Arg	26	ND	ND	rs267607323	DM	Yes	5	Type 2E	Yes	Pathogenic	1/2	·	·
c.3437A>G	p.Tyr1146Cys[d]	26	ND	ND	rs267607326	DM	Yes	5	Type 3[d]	Yes	Pathogenic	1/1	·	·
c.3485_3486 delinsTG	p.Pro1162 Leu	26	ND	ND	NA	ND	Yes	5	Type 3/ Type 2N	Yes	Uncertain significance	2/1	·	·
c.3590C>A	**p.Pro1197Gln**	27	ND	ND	NA	ND	ND	2	Co-inh./2B	No	Benign	1/0		·
c.3686T>C	**p.Val1229Ala**	28	ND	ND	NA	ND	ND	0	Co-inh./2N	No	Benign	0/1		·
c.3692A>C	p.Asn1231Thr	28	0.002190	0.000466	rs61749368	DM	Yes	2	Co-inh./2N	No	Benign	0/1		·
c.3814T>G	p.Cys1272Gly	28	ND	ND	rs61749372	DM	Yes	5	Type 2A	No	Pathogenic	1/0	·	·
c.3815G>T	p.Cys1272Phe	28	ND	ND	rs63524161	DM	Yes	5	Type 2A	No	Pathogenic	1/0	·	·
c.3916C>T	p.Arg1306Trp	28	ND	ND	rs61749384	DM	Yes	4	Type 2B	Yes	Pathogenic	4/4	·	·
c.3943C>T	p.Arg1315Cys	28	ND	ND	rs61749395	DM	Yes	4	Type 2M	No	Pathogenic	1/0	·	·
c.3944G>A	p.Arg1315His	28	0.00003310	0.000116	rs61749396	DM	Yes	4	Type 2M	Yes	Pathogenic	1/1	·	·
c.3946G>A	p.Val1316Met	28	ND	ND	rs61749397	DM	Yes	4	Type 2B	Yes	Pathogenic	1/1	·	·
c.4117G>T	**p.Asp1373Tyr**[d]	28	ND	ND	NA	ND	ND	4	Type 2M[d]	No	Pathogenic	1/0	·	·
c.4120C>T	p.Arg1374Cys	28	0.000008292	ND	rs61750071	DM	Yes	4	Type 2M	No	Pathogenic	2/0	·	·
c.4121G>A	p.Arg1374His	28	ND	ND	rs61750072	DM	Yes	4	Type 2M	No	Pathogenic	4/0	·	·
c.4195C>T	p.Arg1399Cys	28	0.00001660	ND	rs61750077	DM	Yes	4	Type 2M	Yes	Pathogenic	2/2	·	·
c.4257T>G	p.His1419Gln[d]	28	0.0001071	ND	rs569177726	ND	ND	2	Type 2M[d]	No	Benign	1/0	·	·
c.4517C>T	p.Ser1506Leu	28	ND	ND	rs61750100	DM	Yes	5	Type 2A	No	Pathogenic	2/0	·	·
c.4751A>G	p.Tyr1584Cys	28	0.002942	0.003023	rs1800386	DM	Yes	3	Type 1	No	Benign	0/1	·	·
c.4883T>C	p.Ile1628Thr	28	ND	ND	rs61750584	DM	Yes	3	Type 2A	Yes	Pathogenic	4/3	·	·
c.5140G>C	**p.Ala1714Pro**[d]	29	ND	ND	NA	ND	ND	4	Type 1[d]	Yes	Pathogenic	1/1	·	·
c.6890C>T	p.Pro2297Leu	39	0.00005100	ND	rs201372397	DM	Yes	4	Co-inh./2N	No	Likely Pathogenic	1/0		·
c.7400A>C	**p.Gln2467Pro**	43	ND	ND	NA	ND	ND	4	Type 1	No	Pathogenic	1/0	·	·
Nonsense variants														
c.3931C>T	p.Gln1311*	28	ND	ND	NA	DM	Yes	–	Type 3	Yes	Pathogenic	5/1	·	·
c.4666C>T	**p.Gln1556***	28	ND	ND	NA	ND	ND	–	Type 3	Yes	Pathogenic	4/3	·	·

c.7086C>A	**p.Cys2362***	42	ND	ND	NA	ND	ND	–	Type 3	No	Pathogenic	1/0	·	·
Small deletions and duplication														
c.100delT	**p.Arg34Aspfs*49**[d]	3	ND	ND	NA	ND	ND	–	Type 2N [d]	No	Pathogenic	1/0	·	·
c.5414_5415del	**p.Val1805Glyfs*8**	31	ND	ND	NA	ND	ND	–	Type 3	No	Pathogenic	1/0	·	·
c.6699_6702dup	p.Cys2235Argfs*8	38	ND	ND	NA	Yes	Yes	–	Type 3	Yes	Pathogenic	5/1	·	·
Large deletion														
c.5312–104_5455+642del (890bp)		31	ND	ND	NA	ND	ND	–	Type 1	Yes	Pathogenic	1/2		
Synonymous variants								(Supp.Table 3)						
c.1077C>T	p.(=)	9	0.001016	0.000814	rs71582884	ND	Yes	0	Co-inh./2N	No	Likely benign	0/1		·
c.3426T>C	p.(=)	26	ND	ND	NA	ND	Yes	0	Co-inh./3	Yes	Likely benign	2/1	·	·
c.7437G>A	p.(=)[d]	43	ND	ND	rs267607363	Yes	Yes	6 (DSS)	Type 1[d]	No	Pathogenic	1/0	·	·
c.7464C>T	p.(=)[d]	44	ND	ND	NA	ND	Yes	5 (DSS)	Type 2M[d]	No	Likely Pathogenic	1/0		·
c.8175C>T	p.(=)	51	0.000049	ND	NA	ND	ND	0	Co-inh./2M	Yes	Likely benign	1/1		·
Intronic variants		Intron												
c.55+8C>A		2	0.001773	0.000233	rs114713980	ND	Yes	0	Co-inh./2A	No	Likely benign	1/0		·
c.998–46C>T		9	ND	ND	NA	ND	ND	0	Co-inh./2M	No	Likely benign	1/0		·
c.1156+42C>T		10	ND	0.000233	rs376410074	ND	ND	0	Co-inh./2N	No	Likely benign	1/0		·
c.1293+109T>C		11	ND	ND	rs375092486	ND	ND	0	Co-inh./1	No	Likely benign	0/1		·
c.1533+1G>A[d]		13	ND	ND	NA	ND	ND	5 (DSS)	Type3[d]	Yes	Pathogenic	1/1	·	·
c.1534–38G>A		13	ND	ND	NA	ND	ND	0	Co-inh./2M	No	Likely benign	1/0		·
c.2821–57_-2821–56delTCinsA		21	ND	ND	NA	ND	ND	0	Co-inh./3	No	Likely benign	6/1		·
c.2821–102C>T		21	ND	ND	NA	ND	ND	0	Co-inh./2M	No	Likely benign	1/0		·
c.2968–105G>A		23	ND	ND	NA	ND	ND	0	Co-inh./2M	No	Likely benign	1/0		·
c.3380–92_3380–90delTAT		25	ND	ND	NA	ND	ND	0	Co-inh./3	No	Likely benign	0/1		·
c.5170+10C>T		29	0.00364	0.00465	rs61750601	ND	Yes	3 (DSS)	Co-inh./1, 2B	No	Uncertain significance	1/1	·	·
c.5620+33delA		32	ND	ND	NA	ND	ND	1 (DSS)	Co-inh./3	No	Likely benign	2/1		·
c.5621–139A>G		33	ND	ND	NA	ND	ND	0	Co-inh./3	No	Likely benign	0/1		·
c.6799–47G>A		38	ND	ND	NA	ND	ND	0	Co-inh./2M	No	Likely benign	1/0		·
c.7288–68G>A		42	ND	ND	NA	ND	ND	3 (DSS)	Co-inh./2A	No	Uncertain significance	1/0		·
c.7730–4C>G[d]		45	0.0001155	0.000116	rs71581030	ND	ND	1 (ASS)	Type 1[d]	Yes	Likely Pathogenic	1/3	·	·
c.7730–24A>G		45	ND	ND	NA	ND	ND	0	Co-inh./1	No	Likely benign	1/3		·
c.7730–223A>G		45	ND	ND	NA	ND	ND	1 (ASS)	Co-inh./1	No	Likely benign	1/0		·
c.7771–48G>A		46	ND	ND	NA	ND	ND	1 (ASS)	Co-inh./2B	No	Likely benign	1/1		·
c.7771–86G>A		46	ND	ND	NA	ND	ND	3 (ASS)	Co-inh./2M	No	Uncertain significance	1/0		·

[a]Minor allele frequency (MAF) values are from the Exome Aggregation Consortium (ExAC) and the European-American (EA) population from the Exome Variant Server (EVS). Novel variants (**in bold**) are those variants were not reported in the population databases (1000 Genomes, dbSNP database, ExAC, and EVS), Human Gene Mutation Database (HGMD) or VWD-specific databases (ISTH-SSC VWF, LOVD and VWFdb Hemobase). [b]The *in silico* score for missense variants was calculated using PROVEAN, SIFT, PolyPhen-2, MutationAssessor, and MutationTaster (Suppl. Table 2) and were considered deleterious when ≥ 3. [c]The *in silico* score for predicting potential splice site variants was calculated using HSF, MaxEntScan, NNSPLICE, NetGene2 and Splice View (Suppl. Table 3) and was considered deleterious for a score ≥ 3 (intronic variants) or score≥ 4 (exonic variants). [d]VWD subtype defined with the other variant in the compound heterozygous state. Co-inh./ designates the variants that were co-inherited with other variants and had an unknown functional effect on VWD subtype. Variants detected by Sanger and NGS sequencing are marked by a **dot.** Variants that initially were not screened using Sanger sequencing are marked by a grey cell. Prop. *propositus*; NA, not available; ND, no data; ASS, acceptor splice site; DSS, donor splice site.

Suppl. Table 2 - Summary of *in silico* analysis (PROVEAN, SIFT, Polyphen-2, Mutationassessor and Mutation Taster) for missense variants.

Nucleotide Change	Amino Acid Change	Exon	PROVEAN[a]		SIFT[b]		PolyPhen_2[c]		MutAss[d]		MutTaster[e]		
			Score	prediction	Score	prediction	Score	prediction	Score	prediction	Score	prediction	*In silico* score
c.440A>G	**p.Gln147Arg**	5	-1.227	Neutral	0.059	Tolerated	0.8	Possibly damaging	2.135	Medium	0.99	Disease causing	3
c.1027G>T	p.Val343Leu	9	-0.661	Neutral	0.04	Deleterious	0.992	Probably damaging	3.115	High	1	disease causing	4
c.1446C>G	p.Ile482Met	13	-2.107	Neutral	0.001	Deleterious	0.991	Probably damaging	2.99	Medium	1	Disease causing	4
c.1892C>T	p.Ala631Val	15	-1.689	Neutral	0.29	Tolerated	0.594	Possibly damaging	2.995	Medium	0.98	Disease causing	3
c.2446C>T	p.Arg816Trp	19	-4.378	Deleterious	0.001	Deleterious	1.0	Probably damaging	2.075	Medium	1	Disease causing	5
c.2451T>A	p.His817Gln	19	-2.144	Neutral	0.03	Deleterious	0.09	Benign	1.69	Low	1	Disease causing	2
c.2561G>A	p.Arg854Gln	20	-3.578	Deleterious	0.08	Tolerated	0.994	Probably damaging	1.765	Low	1	Disease causing	3
c.2637C>A	**p.Asp879Glu**	20	-3.841	Deleterious	0.0	Deleterious	0.982	Probably damaging	4.305	High	1	Disease causing	5
c.3388T>C	p.Cys1130Arg	26	-11.773	Deleterious	0.0	Deleterious	0.998	Probably damaging	3.825	High	1	Disease causing	5
c.3437A>G	p.Tyr1146Cys	26	-8.823	Deleterious	0.0	Deleterious	1.0	Probably damaging	4.18	High	1	Disease causing	5
c.3485_3486 delinsTG	p.Pro1162 Leu	26	-4.905	Deleterious	0.04	Deleterious	0.986	Probably damaging	2.5	Medium	1	Disease causing	5
c.3590C>A	**p.Pro1197Gln**	27	-1.535	Neutral	0.3	Tolerated	0.972	Probably damaging	1.75	Low	1	Disease causing	2
c.3686T>C	**p.Val1229Ala**	28	1.051	Neutral	0.34	Tolerated	0.0	Benign	0.0	Low	1	polymorphism	0
c.3692A>C	p.Asn1231Thr	28	-1.121	Neutral	0.39	Tolerated	0.651	Possibly damaging	1.73	Low	0.98	Disease causing	2
c.3814T>G	p.Cys1272Gly	28	-4.723	Deleterious	0.0	Deleterious	0.999	Probably damaging	3.32	Medium	1	Disease causing	5
c.3815G>T	p.Cys1272Phe	28	-3.564	Deleterious	0.0	Deleterious	0.998	Probably damaging	3.32	Medium	1	Disease causing	5
c.3916C>T	p.Arg1306Trp	28	-1.585	Neutral	0.0	Deleterious	0.981	Probably damaging	3.795	High	0.93	Disease causing	4
c.3943C>T	p.Arg1315Cys	28	-1.348	Neutral	0.0	Deleterious	1.0	Probably damaging	4.01	High	1	Disease causing	4
c.3944G>A	p.Arg1315His	28	-2.039	Neutral	0.0	Deleterious	0.799	Possibly damaging	3.66	High	1	Disease causing	4
c.3946G>A	p.Val1316Met	28	-1.065	Neutral	0.01	Deleterious	0.999	Probably damaging	3.955	High	1	Disease causing	4
c.4117G>T	**p.Asp1373Tyr**	28	-2.524	Deleterious	0.01	Deleterious	0.998	Probably damaging	3.455	High	0.98	Polymorphism	4
c.4120C>T	p.Arg1374Cys	28	-0.137	Neutral	0.0	Deleterious	0.999	Probably damaging	4.015	High	1	Disease causing	4
c.4121G>A	p.Arg1374His	28	-1.383	Neutral	0.0	Deleterious	0.998	Probably damaging	4.015	High	1	Disease causing	4
c.4195C>T	p.Arg1399Cys	28	-0.342	Neutral	0.01	Deleterious	0.998	Probably damaging	3.795	High	1	Disease causing	4
c.4257T>G	p.His1419Gln	28	-1.673	Neutral	0.09	Tolerated	0.819	Possibly damaging	2.955	Medium	1	Polymorphism	2
c.4517C>T	p.Ser1506Leu	28	-3.087	Deleterious	0.01	Deleterious	1.0	Probably damaging	4.4	High	1	Disease causing	5
c.4751A>G	p.Tyr1584Cys	28	-0.394	Neutral	0.02	Deleterious	0.981	Probably damaging	4.125	High	1	Polymorphism	3
c.4883T>C	p.Ile1628Thr	28	-0.879	Neutral	0.0	Deleterious	0.229	Benign	3.955	High	1	Disease causing	3
c.5140G>C	**p.Ala1714Pro**	29	-0.972	Neutral	0.03	Deleterious	0.974	Probably damaging	2.51	Medium	0.91	Disease causing	4
c.6890C>T	p.Pro2297Leu	39	-2.620	Deleterious	0.01	Deleterious	0.261	Benign	2.505	Medium	1	Disease causing	4
c.7400A>C	**p.Gln2467Pro**	43	-1.193	Neutral	0.02	Deleterious	0.898	Possibly damaging	2.015	Medium	1	Disease causing	4

[a]Provean uses the alignment and the measurement of the similarity between the variant sequence and the protein sequence homolog. The score threshold is set at -2.5 for binary classification (i. e., deleterious *vs* neutral). [b]SIFT is used to predict the effect of sequence changes on the protein function and is based on a homology search and the physical properties of amino acids – scores range from 0 to 1. An amino acid substitution is predicted to be damaging if the score is <=0.05 and is tolerated if the score is >0.05.

[c]PolyPhen is a tool that predicts the possible effect of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. Scores between 0.909 and 1 denote a probably damaging variant, a score between 0.447 and 0.908 indicates a possibly damaging variant, and a score below 0.446 is a benign variant.

[d]MutationAssessor, The score predicts the functional impact of amino acid substitutions, with high scores denoting a more severe impact.

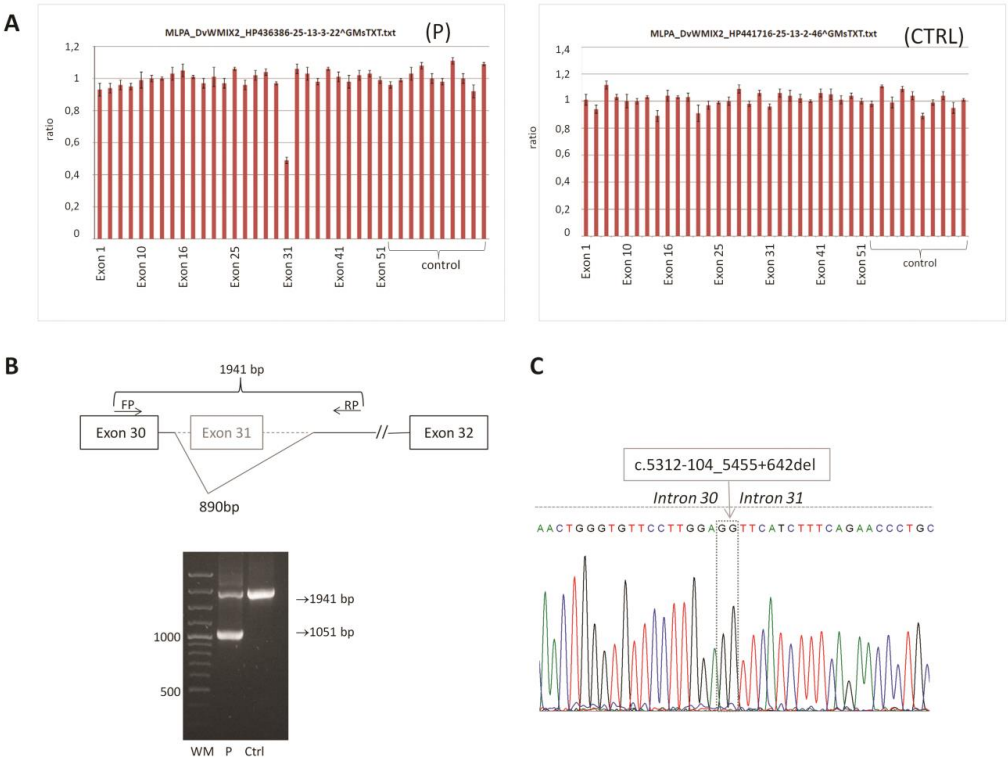
[e]MutationTaster uses a Bayes classifier to calculate the probability of whether the alteration in the sequence is a disease mutation or a harmless polymorphism. A probability close to 1 indicates a high certainty in the prediction. The *in silico* scores were considered deleterious when ≥ 3. Novel variants are marked in **bold**.

Suppl. Table 3. Summary of potential splice site variants and predicted impact on RNA processing. A higher score implies greater potential for splice site.

Nucleotide Change	HSF score native-mutated	MaxEntScan score native-mutated	NNSplice score native-mutated	NetGene2 score native-mutated	Splice View score native-mutated	MutPred Splice	*In silico* score	Comments
Intronic variants								
c.1533+1G>A	90.71–63.87 Broken DSS	10.29–2.1	0.98 - native DSS destroyed	1.00 - native DSS destroyed	90 - native DSS destroyed	NA	5	Heterozygous in *trans* with c.3437A>G, p.Tyr1146Cys
c.5170+10C>T	52.66–79.49 New DSS	No difference	0.84–0.97	No difference	Not predited – New DSS -78	NA	3	Heterozygous in *trans* with c.5140G>C, p.Ala1714Pro; Heterozygous in *trans* with c.3946G>A, p.Val1316Met
c.5620+33delA	68.43–9.69 Broken DSS	No difference	No difference	No difference	No difference	NA	1	Homozygous with c.3485_3486delinsTG
c.7288–68G>A	59.3–67.64 New DSS	2.89–6.3	No difference	No difference	Not predited – New DSS -75	NA	3	Heterozygous in *trans* with c.3815G>T, p.Cys1272Phe
c.7730–4C>G	36.76–65.71 New ASS	No difference	No difference	No difference	No difference	NA	1	Homozygous with c.1892C>T, p.Ala621Val
c.7730–223A>G	41.08–70.02 New ASS	No difference	No difference	No difference	No difference	NA	1	Heterozygous with c.6699_6702dupAGGC, p.Cys2235Argfs*8 and c.7437G>A, p. (=)
c.7771–48G>A	40.85–69.79 New ASS	No difference	No difference	No difference	No difference	NA	1	Heterozygous in *trans* with c.3916C>T, p. Arg1306Trp; Heterozygous in *trans* with c.3931C>T, p.Gln1311*
c.7771–86G>A	36.63–65.68 New ASS	3.82–8.73	Not predited - 0.98 New DSS	No difference	No difference	NA	3	Heterozygous with c.4121G>A, p.Arg1374His
Exonic variants								
c.2451T>A	40.74–69.68 New ASS	No difference	No difference	No difference	No difference	0.24 (SNV)	1	Heterozygous in *trans* with c.3485_3486delinsTG
c.2561G>A	50.04–78.99 New ASS	No difference	No difference	No difference	No difference	0.26 (SNV)	1	Homozygous
c.2637C>A	52.25–81.19 New ASS	-2.76–5.27	No difference	No difference	Not predited – New ASS -82	0.31 (SNV)	3	Heterozygous in *trans* with c.2561G>A p.Arg854Gln
c.3437A>G	41.44–68.27 New DSS	No difference	No difference	No difference	No difference	0.19 (SNV)	1	Heterozygous in *trans* with c.1533+1G>A
c.4257T>G	57.71–86.65 New ASS	-4.84–3.75	No difference	No difference	No difference	0.15 (SNV)	2	Heterozygous with c.1027G>T, p.Val343Leu; c.1446C>G, p.Ile482Met, c.4117G>T, p.Asp1373Tyr
c.7437G>A	92.11–81.53 Broken DSS	3.52–7.03	0.98–0.61	0.55–0.92	Not predited – New DSS -79	0.9 (SAV); Loss of natural 5' SS; (P < 0.000001)	6	Heterozygous in *trans* with c.6699_6702dupAGGC, p.Cys2235Argfs*8
c.7464C>T	No difference	2.72–10.47	Not predited - 0.99 New DSS	Not predited - DSS -0.63	Not predited – New DSS -94	0.97 (SAV) Cryptic 5' SS	5	Heterozygous in *trans* with c.3944G>A, p.Arg1315His

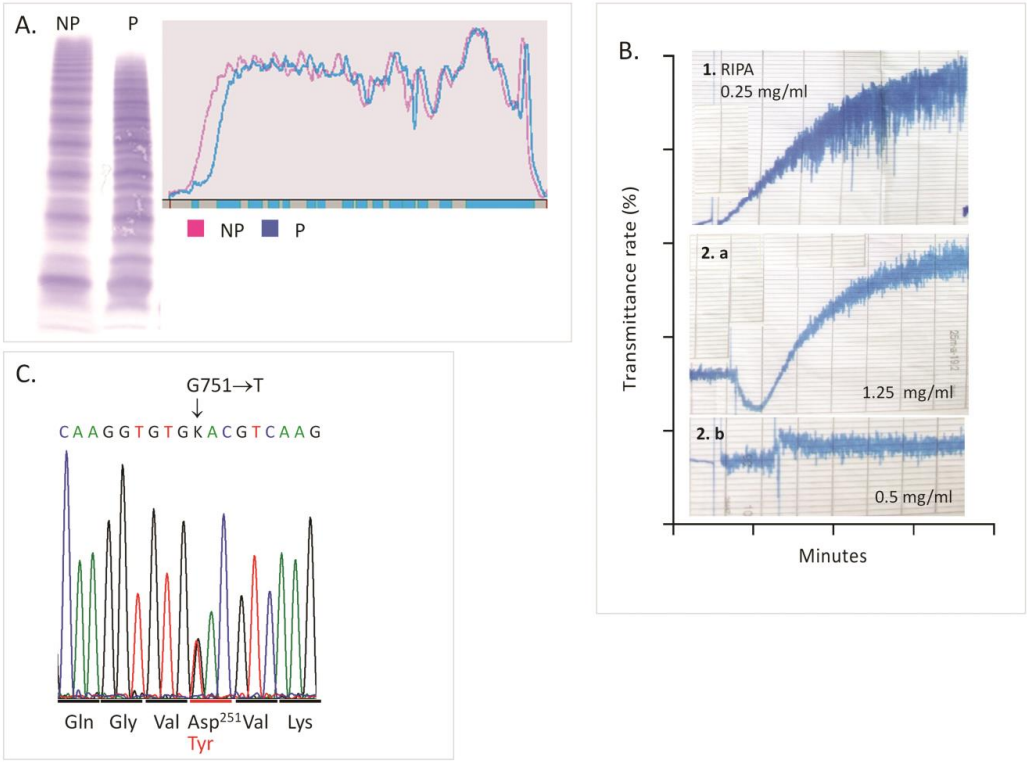
ASS: acceptor splice site; DSS: donor splice site; NA – not applicable; HSF score (0–100); NNSplice score (0–1); NetGene2 score (0–1); MaxEntScan score: maximum entropy score; Splice View score (0–100). MutPred Splice - Not predicted to disrupt splicing, Splice Neutral Variant (SNV), (general score <0.6); Confident calls of splicing variants - predicted Splice Affected Variant (SAV), (general score >=0.6). *In silico* scores were considered deleterious when ≥ 3 for intronic variants and when ≥ 4 for exonic variants. Novel variants are marked in **bold**.

Suppl. Figure 1



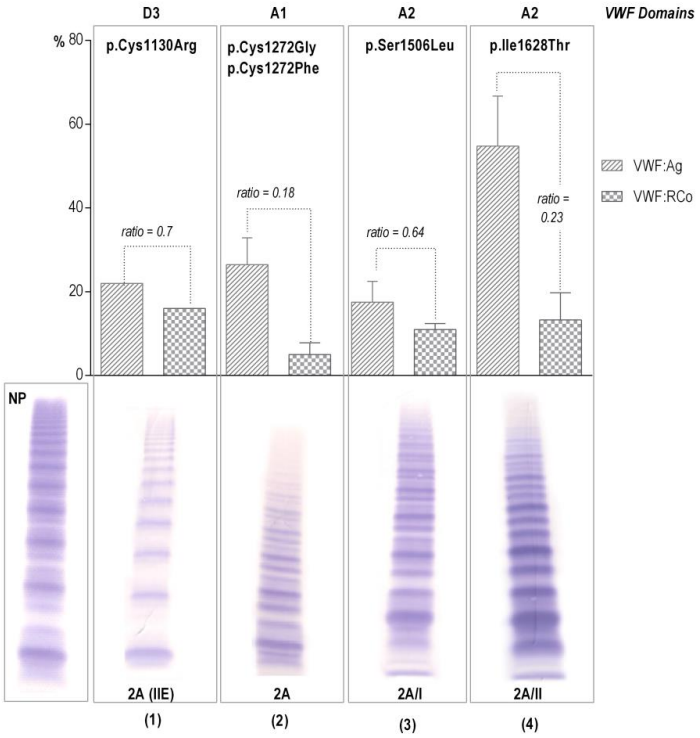
Suppl. Figure 1: Detection of the *VWF* exon 31 deletion breakpoints. (A) Detection of exon 31 deletion following dosage analysis using MLPA in the proband (P) and a healthy control (CTRL). (B) The PCR designed to detect a deletion (890 bp) shows the heterozygous state in the proband (P). FP - forward primer; RP, reverse primer and (C) deletion breakpoints identification (c.5312–104_5455+642del) by Sanger sequencing.

Suppl. Figure 2



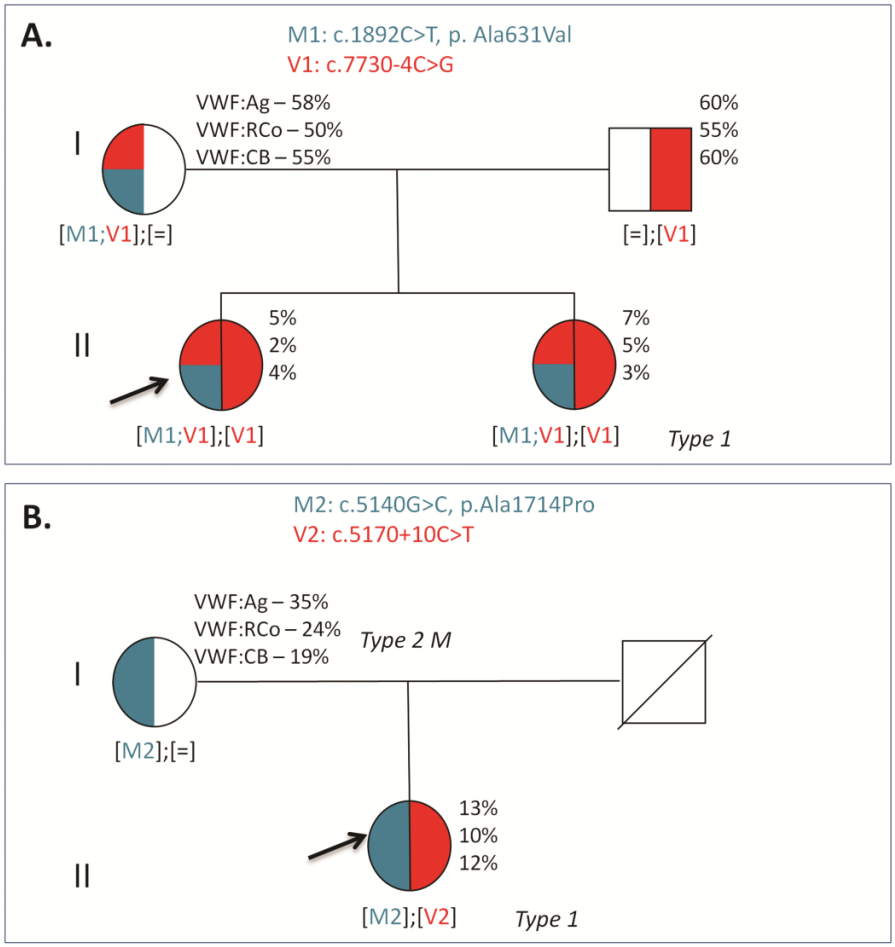
Suppl. Figure 2: Laboratory data of the PT-VWD patient. (A) Plasma VWF multimer analysis using 1.6 % SDS-agarose gel electrophoresis and densitometry showing partial loss of the HMW multimers. NP = Normal plasma; P = Patient plasma; (B) 1. RIPA in the platelet-rich plasma of the patient showing aggregation with a low concentration of ristocetin (0.25 mg/ml); 2. RIPA mixing studies: a) patient plasma/control platelets showing a normal aggregation response at 1.25 mg/ml ristocetin and b) showing the lack of an aggregation response at 0.5 mg/ml; (C) *GP1BA* sequencing showing heterozygosity for c.751G>T, p.Asp251Tyr.

Suppl. Figure 3



Suppl. Figure 3: Five missense mutations identified in probands with VWD 2A structural defects, their VWF:RCo/VWF:Ag ratios and the multimer VWF patterns. The VWF:RCo and VWF:Ag levels shows that p.Cys1130Arg and p.Ser1506Leu had a disproportionate VWF:RCo/VWF:Ag ratio that was close to 0.7; p.Cys1272Gly/Phe and p.Ile1628Thyr had lower values that were indicative of a higher extent of proteolysis. The VWF multimers analysed by SDS-agarose gel electrophoresis displayed a characteristic pattern in each mutation: (1) – multimerization defect (p.Cys1130Arg in domain D3) with the absence of large multimers and no triplet structure; (2) – change in protein folding (p.Cys1272Gly/Phe in loop of A1 domain) with a loss of large and intermediate multimers; (3) – intracellular proteolysis of large VWF multimers, (p.Ser1506Leu) with a loss of large multimers; (4) – hypersensitivity to ADAMTS13 (p.Ile1628Thr in A2 domain) with the absence of large VWF multimers and increased triplet structure. NP = Normal plasma

Suppl. Figure 4



Suppl. Figure 4 - Pedigree and laboratory features associated with c.7730–4C>G and c.5170+10C>T variants. The study of the parents of these probands allowed the tracing of the mutation inheritance on each allele. The variants c.7730–4C>G in the homozygous state (A) and c.5170+10C>T (B) in the compound heterozygous state appear to have contributed to the changes in VWD phenotype in each proband. The arrow indicates the *propositus*; M, mutation; V, variant; and [=],'no change' in the other allele.

Suppl. References

1. Protein Variation Effect Analyzer (PROVEAN). Available from: http://provean.jcvi.org/index.php.
2. Sorting Intolerant from Tolerant (SIFT). Available from: http://sift.jcvi.org.
3. PolyPhen-2 prediction of functional effects of human nsSNPs. Available from: http://genetics.bwh.harvard.edu/pph2.
4. Functional impact of protein mutations. Available from: http://mutationassessor.org/v1.
5. Mutation taster. Available from: www.mutationtaster.org.
6. MutPred. Available from: http://mutpred.mutdb.org.
7. NetGene2 . Available from: (http://www.cbs.dtu.dk/services/NetGene2/).
8. NNSplice. Available from: (http://www.fruitfly.org/seq_tools/splice.html).
9. MaxEntScan. Available from: (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html).
10. SpliceView. Available from: (http://zeus2. itb.cnr.it/~webgene/wwwspliceview.html).
11. Human Splicing Finder. Available from: (http://www.umd.be/HSF/).

ARTICLE 4

Títol: "Desxifrant l'efecte de mutacions silencioses, intròniques i *missense* en l'*splicing* del *VWF*: aportació de la NGS a l'estudi de l'mRNA"

Resum

Els grans estudis de cohorts de pacients amb la malaltia de von Willebrand (VWD), inclosos els registres espanyol i portuguès, van portar a la identificació de >250 mutacions diferents. Determinar l'efecte patogènic d'un mutació potencial d'*splicing* (PSSM) sobre l'mRNA del gen del factor von Willebrand (*VWF*) és tot un repte. Aquest estudi pretén aclarir els efectes reals de 18 PSSM en el processament de l'mRNA del *VWF*, investigar la contribució de la seqüenciació massiva (NGS) en l'estudi de l'mRNA *in vivo* en la VWD i comparar-ne els resultats amb les anàlisi de predicció *in silico*. L'RNA extret de plaquetes de pacients i leucòcits va ser amplificat per RT-PCR i seqüenciat mitjançant les tècniques de *Sanger* i NGS. Vuit mutacions van afectar l'*splicing* del *VWF*: c.1533+1G> A, c.5664+2T> C i c.546G>A (p.=) van provocar el salt d'un exó; c.3223-7_3236dup i c.7082-2A>G van provocar l'activació de llocs críptics d'*splicing*; c.3379+1G>A i c.7473G>A (p.=) van demostrar ambdós mecanismes moleculars patogènics simultàniament; i la mutació *missense* p.Cys370Tyr va generar dos transcrits aberrants. Cal destacar que l'efecte total de 3 mutacions només va ser proporcionat per la NGS a causa de la baixa expressió dels transcrits aberrants. En les 10 mutacions restants, no es va identificar cap efecte mitjançant els experiments realitzats. No obstant, les troballes diferencials obtingudes en les plaquetes i els leucòcits van proporcionar proves substancials que 4 d'aquestes mutacions tindrien un efecte en els nivells del VWF. Aquest és el primer article que utilitza la tecnologia NGS per desxifrar els efectes de les mutacions del *VWF* en el processament de l'mRNA. Les nostres dades posen de manifest la importància d'estudiar l'efecte de les mutacions sinònimes i *missense* en l'*splicing* del *VWF* per millorar el coneixement actual dels mecanismes moleculars que es troben darrere de la VWD.

Referència

Nina Borràs, Gerard Orriols, Javier Batlle, Almudena Pérez-Rodríguez, Teresa Fidalgo, Patricia Martinho, María Fernanda López-Fernández, Ángela Rodríguez-Trillo, Esther Lourés, Rafael Parra, Carme Altisent, Francisco Vidal, and Irene Corrales; on behalf of the members of the PCM-EVW-ES group.

Unraveling the effect of silent, intronic, and missense mutations on vwf splicing: contribution of NGS to mRNA study.

Haematologica (en revisió).

Haematologica
HAEMATOL/2018/203166
Version 2

Unraveling the effect of silent, intronic and missense mutations on VWF splicing: contribution of next generation sequencing in the study of mRNA.

Nina Borràs, Gerard Orriols, Francisco Javier Batlle, Almudena Pérez-Rodríguez, Teresa Fidalgo, Patricia Martinho, María Fernanda López-Fernández, Ángela Rodríguez-Trillo, María Esther Lourés, Rafael Parra, Carmen Altisent, Ana Rosa Cid, Santiago Bonanad, Noelia Cabrera, Andrés Moret, María Eva Mingot-Castellano, Nira Navarro, Rocío Pérez-Montes, Sally Marcellini, Ana Moreto, Sonia Herrero, Inmaculada Soto, Núria Fernández-Mosteirín, Víctor Jiménez-Yuste, Nieves Alonso, Aurora de Andrés-Jacob, Emilia Fontanes, Rosa Campos, María José Paloma, Nuria Bermejo, Ruben Berrueco, José Mateo, Karmele Arribalzaga, Pascual Marco, Ángeles Palomo, Nerea Castro Quismondo, Belén Iñigo, María del Mar Nieto, Rosa Vidal, María Paz Martínez, Reyes Aguinaco, Jesús María Tenorio, María Ferreiro, Javier García-Frade, Ana María Rodríguez-Huerta, Jorge Cuesta, Ramón Rodríguez-González, Faustino García-Candel, Manuela Dobón, Carlos Aguilar, Francisco Vidal, and Irene Corrales

Disclosures: None.

Contributions: Contribution: This study derived from PCM-EVW-ES registry. I.C. and F.V coordinated the study; N.B and G.O selected the patients included in the study; J.B, A.P.R, MF.L.F, T.F, P.M, R.P and C.A obtained patient's sample; N.B, G.O, F.V. and I.C. analyzed and interpreted results; N.B, F.V. and I.C. wrote the paper; and all the authors that participated in the PCM-EVW-ES registry checked the final version of the manuscript.

Unraveling the effect of silent, intronic and missense mutations on *VWF* splicing: contribution of next generation sequencing in the study of mRNA

Nina Borràs,[1,2] Gerard Orriols,[1] Javier Batlle,[3] Almudena Pérez-Rodríguez,[3] Teresa Fidalgo,[4] Patricia Martinho,[4] María Fernanda López-Fernández,[3] Ángela Rodríguez-Trillo,[3] Esther Lourés,[3] Rafael Parra,[1,5] Carme Altisent,[2] Francisco Vidal,[1,2,6] and Irene Corrales[1,2]; on behalf of the members of the PCM-EVW-ES group.

[1]Banc de Sang i Teixits, Barcelona, Spain; [2]Vall d'Hebron Research Institute, Universitat Autònoma de Barcelona (VHIR-UAB), Barcelona, Spain; [3]Complexo Hospitalario Universitario A Coruña, INIBIC, A Coruña, Spain; [4]Centro Hospitalar e Universitário de Coimbra (CHUC) – Serviço de Hematologia Clínica, Unidade de Trombose e Hemostase, Hospital Pediátrico de Coimbra, Coimbra, Portugal; [5]Hospital Universitari Vall d'Hebron, Barcelona, Spain; and [6]CIBER de Enfermedades Cardiovasculares.

Statement of equal authors' contribution: IC and FV contributed equally to this study.

Running head: Effect of VWF mutations on mRNA splicing

Contact information for correspondence:

Irene Corrales Insa, PhD. Banc de Sang i Teixits, Barcelona, Spain. E-mail: icorrales@bst.cat

Francisco Vidal Pérez, PhD. Banc de Sang i Teixits, Barcelona, Spain. E-mail: fvidal@bst.cat

ABSTRACT

Large studies in von Willebrand disease patients, including Spanish and Portuguese registries, led to identification of >250 different mutations. It is a challenge to determine the pathogenic effect of potential splice site mutations on *VWF* mRNA. This study aimed to elucidate the true effects of 18 mutations on *VWF* mRNA processing, investigate the contribution of next-generation sequencing to *in vivo* mRNA study in von Willebrand disease, and compare the findings with *in silico* prediction. RNA extracted from patient platelets and leukocytes was amplified by RT-PCR and sequenced using Sanger and next generation sequencing techniques. Eight mutations affected *VWF* splicing: c.1533+1G>A, c.5664+2T>C and c.546G>A (p.=) prompted exon skipping; c.3223-7_3236dup and c.7082-2A>G resulted in activation of cryptic sites; c.3379+1G>A and c.7473G>A (p.=) demonstrated both molecular pathogenic mechanisms simultaneously; and the p.Cys370Tyr missense mutation generated two aberrant transcripts. Of note, the complete effect of 3 mutations was provided by next generation sequencing alone because of low expression of the aberrant transcripts. In the remaining 10 mutations, no effect was elucidated in the experiments. However, the differential findings obtained in platelets and leukocytes provided substantial evidence that 4 of these would have an effect on VWF levels. In this first report using next generation sequencing technology to unravel the effects of *VWF* mutations on splicing, the technique yielded valuable information. Our data bring to light the importance of studying the effect of synonymous and missense mutations on *VWF* splicing to improve the current knowledge of the molecular mechanisms behind von Willebrand disease.

ClinicalTrials.gov identifier: NCT02869074

INTRODUCTION

Von Willebrand disease (VWD), the most common congenital bleeding disorder, is caused by a genetic defect in the von Willebrand factor gene (*VWF*).[1] *VWF* mutational analysis can be valuable for diagnosing and investigating the molecular etiology of VWD, as was seen in the Spanish (PCM-EVW-ES project) and Portuguese cohort of VWD patients.[2-4] One interesting challenge in this condition is to elucidate the pathogenic mechanism of *VWF* mutations. *In silico* analysis is considered a suitable supporting tool to predict the pathogenicity of the variants identified.[5, 6] However, functional studies remain essential to unequivocally determine their deleterious effect.[7]

Functional studies can be performed by analyzing the potential effect of splice site mutations (PSSM) in RNA. In addition to splice site consensus sequence mutations, deep intronic, missense, and synonymous mutations can also disturb splicing. Along this line, 25% of synonymous mutations positioned at exon-intron boundaries result in altered splicing, which, in itself, can cause disease, modify the severity of the disease phenotype, or be linked with disease susceptibility.[8] In *VWF*, heterozygotes for PSSM may be associated with mild forms of VWD type 1 or be phenotypically silent, but when 2 such mutations are found in different alleles, the phenotype is associated with VWD type 3.[9]

Almost all related studies of splicing effects have examined peripheral blood platelets, as *VWF* is exclusively expressed in these cells and endothelial cells.[10, 11] Platelets are anucleated, but they contain small amounts of translationally active megakaryocytic mRNA.[12,13] In contrast, the amount of mRNA obtained from leukocytes is higher and contains mRNA transcripts for genes that are not normally expressed in these cells, known as ‘‘ectopic transcripts’’. Their analysis has been used to investigate mutations in several inherited disorders,[14,15] as they facilitate the study of mRNA of those genes expressed in hard to reach tissues.[16] PSSM can affect mRNA reorganization and introduce premature termination codons (PTCs) into open reading frames, a common cause of genetic disorders. Most nonsense transcripts are recognized and degraded by nonsense-mediated mRNA decay (NMD),[17] a degradation pathway to control synthesis

of truncated proteins.[18] The efficiency of NMD varies between cell types; hence, the use of RNA from platelets and leukocytes for *in vivo* study of *VWF* PSSM offers complementary results, particularly when NMD occurs in the allele carrying the mutation in platelets, as we reported.[7]

Since its development, next-generation sequencing (NGS) has been increasingly used in molecular genetics to identify mutations causing disease. However, few groups have explored its potential for analyzing splicing variants following RT-PCR.[19, 20] In this new scenario, the procedure we previously described to analyze the effects of PSSM in *VWF*[7] has been optimized and adapted to an NGS-based technique to investigate its value in this field. Our main objective was to elucidate the true effects of 18 selected mutations (intronic, synonymous, delins, and missense) on mRNA processing and their genotype/phenotype correspondence by analysis of leukocytes and platelets from clinical samples. Finally, the *in vivo* effects of the mutations were compared with the *in silico* predictions.

METHODS

Patients

We studied 15 patients diagnosed with different types of VWD, 5 from Complexo Hospitalario Universitario A Coruña, 8 from Vall d'Hebron University Hospital (HUVH), and 2 from Centro Hospitalar e Universitário de Coimbra. Samples from 4 healthy individuals were used as controls. The study was performed according to the guidelines of the Declaration of Helsinki and was approved by the local Research Ethics Committee. All participants provided written informed consent.

Splice site prediction software

The predicted impact of potential splice site mutations was analyzed with NetGene2[21] and the splicing prediction module of Alamut Visual v.2.6.1 software (Interactive Biosoftware, Rouen, France), which integrates data from 3 methods: Splice Site

Prediction by Neural Network (NNSplice), MaxEntScan, and Human Splicing Finder (HSF).

Platelet and leukocyte separation and RNA isolation

Leukocyte and platelet RNA from patients and controls was isolated from 10 mL of peripheral blood collected in EDTA tubes, as previously described.[7]

VWF mRNA amplification

After RNA isolation, cDNA was synthesized using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's recommendations. The region including the mutation was amplified by Platinum Taq DNA Polymerase (Thermo Fisher Scientific) in leukocyte and platelet cDNA (Online Supplementary Methods and Table S1). PCR products were separated on 1% agarose gel and visualized by SYBR Safe DNA Gel Stain (Thermo Fisher Scientific).

Sanger sequencing and analysis

PCR products were sequenced as previously described.[22] However, multiple-band PCR products were previously agarose-purified using the MiniElute Gel Extraction kit (Qiagen). The sequences obtained were assembled and aligned against the consensus wild-type (WT) *VWF* mRNA sequence (GenBank NM_000552) using SeqScape v2.7 software (Thermo Fisher Scientific).

Next-generation sequencing and analysis

PCR amplicons obtained per patient were equimolarly mixed in a single tube in a total amount of 250 ng. Subsequently, the libraries were fragmented and the adapter and barcodes were ligated using the NGSgo protocol (GenDX, Utrecht, Netherlands) following the manufacturer's recommendations. Resulting libraries were combined and sequenced on a MiSeq platform (Illumina, San Diego, CA).

After sequencing, barcoded sequences were demultiplexed and analyzed individually. The paired sequence files (fastq format) were used as input for analysis with the CLC Genomic Workbench v.11 software (Qiagen, Aarhus, Denmark) (Online Supplementary Methods and Figure S1-S2).

RESULTS

An in-depth study was performed in PCM-EVW-ES,[3] the Portuguese cohort,[4] and HUVH patients to select previously undescribed *VWF* mutations and mutations with an unknown or controversial pathogenic mechanism. Eighteen mutations (15 patients) were selected: 8 intronic (4 in canonical, GT and AG, splice site sequences), 5 synonymous, 2 missense, and 3 delins. The patients' phenotypic and molecular data are summarized in Table 1 and Online Supplementary Table S2. Total mRNA was obtained from platelets and leukocytes of all patients, with the exception of patients UMP08 and UMP14, in whom platelet RNA isolation failed due to blood lysis. All mutations were analyzed by both Sanger sequencing and NGS, and results were compared to the predictions generated by *in silico* analysis (Online Supplementary Table S3).

Mutations in canonical splice site sequences

The **c.1533+1G>A** mutation (intron 13) was identified in a type 3 VWD carrier (UMP01). The exon 11-15 region was analyzed with a specific primer pair to avoid amplification of a prevalent alternative-splicing product (skipping of exons 14 and 15) in *VWF* from leukocytes.[7] Whereas only the expected PCR product was observed in platelets, leukocyte amplification resulted in 4 PCR bands. Sanger analysis of leukocytes showed 2 mRNA aberrant transcripts: 1) lacking exon 13; and 2) lacking exon 13 and 14 (Online Supplementary Figure S3). By means of NGS it was determined that 19% were transcripts without exon 13, 45% transcripts without exons 13 or 14, and an additional aberrant transcript without exon 14 was detected in 8% transcripts (Online Supplementary Table S4). By both techniques, no effect was visible in platelets due to NMD.

Table 1. *Laboratory and molecular data of VWD patients*

Patients code	VWD	VWF: Ag	VWF: RCo	FVIII :C	NT change	AA change	Exon	Intron	Domain
UMP01	3C	70	70.5	111	**c.1533+1G>A**	-	-	13	intronic
UMP02	1H	36	49	101	**c.3379+1G>A** *	-	-	25	intronic
					c.-2627C>T *	-	-	-	upstream
UMP03	1	26	21.2	64.8	**c.5664+2T>C** *	-	-	33	intronic
					c.7220T>C *	p.Leu2407Pro	42	-	C2
UMP04	1	50	52	47	**c.7081+6G>T**	-	-	41	intronic
UMP05	1	27	32	72	**c.7730-56C>T**	-	-	45	intronic
UMP06	2A/ 2M	12	4	24	c.4120C>T *	p.Arg1374Cys	28		A1
					c.7730-4C>G *	-	-	45	intronic
UMP07	1	37	28	60	**c.8254-5T>G**	-	-	51	intronic
UMP08	1	42.4	41.7	73.4	**c.546G>A** †	p.Ser182	6	-	D1
					c.4866C>T †	p.Asp1622	28	-	A1-A3
UMP09 ‡	1	51	49.8	71	**c.3291C>T**	p.Cys1097	25	-	D3
UMP10	3C	46.4	45	105	**c.1109G>A**	p.Cys370Tyr	9	-	D1
UMP11	1	31	24	57	**c.3223-7_3236dup**	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	25	-	D3
UMP12	3	3.7	5	6.7	**c.449T>C** †	p.Leu150Pro	5	-	D1
					c.7082-2A>G †	-	-	41	intronic
UMP13	2A	117	6.5	90	**c.3426T>C** §	p.Cys1142	26	-	D3
					c.3485_3486delinsTG §	p.Pro1162Leu	26	-	D3
					c.8318G>C †	p.Cys2773Ser	52	-	CK
UMP14	1	6	6	4,5	**c.6699_6702dup** †	p.Cys2235ArgTer8	38	-	D4
					c.7437G>A †	p.Ser2479	43	-	C3
UMP15||	3	2	<1	5	c.546G>A §	p.Ser182	6	-	D1
					c.7082-2A>G §	-	-	41	intronic
					c.8155+3G>C †	-	-	50	intronic

*All mutations were identified in heterozygous state. In bold, mutations selected to study their effect on VWF mRNA. AA indicates amino acid; NT, nucleotide and; VWD, von Willebrand disease. *The allelic phase (cis/trans) of mutations could not be determined since no informative relatives are available. †Mutations in trans. ‡Patient with an additional mutation in the F8 gene. §Mutations in cis. ||Patient previously studied at mRNA level by Corrales et al.*[7]

The **c.3379+1G>A** mutation (intron 25) was identified in a type 1H VWD patient (UMP02). We designed 2 new primers to analyze the exon 22-26 region. Amplification of leukocyte cDNA yielded 2 bands, one with the expected size and a smaller band, whereas in platelets only the expected PCR product was observed (Online Supplementary Figure S4). Sanger analysis of leukocytes showed that c.3379+1G>A caused exon 25 skipping, leading to a frameshift at position 1075 and adding 88 aberrant amino acids before a PTC was encountered (p.Pro1075ValfsTer88). On NGS, however, 2 aberrant transcripts were detected: the major one, which lacked exon 25, was detected in 43% of all reads, and the minor one, which resulted from activation of

a cryptic DSS 31 nucleotides upstream from the WT-DSS, was found in 1.4% of reads (p.Pro1117ValfsTer88) (Table 2 and Online Supplementary Table S4). In platelets, however, the mutated allele had undergone NMD, which precluded observation of aberrant transcripts.

The **c.5664+2T>C** mutation was identified in a type 1 VWD patient (UMP03), *in trans* with the p.Leu2407Pro mutation. To analyze the PSSM in intron 33, the exon 30-35 region was amplified. The PCR product had the expected size in platelets, whereas an additional smaller one was observed in leukocytes (Figure 1). Sanger sequencing of the leukocyte PCR product confirmed exon 33 skipping (p.Gly1874AlafsTer32) (Table 2). NGS detected 2 aberrant transcripts: the major one (37% of reads) showed exon 33 skipping, and the minor (2% of reads) showed skipping of exons 33 and 34. Moreover, analysis of the p.Leu2407Pro mutation confirmed that NMD of the allele carrying the c.5664+2T>C mutation had occurred in platelets.

Candidate intronic mutations

The **c.7081+6G>T** mutation, identified in a type 1 VWD patient (UMP04), generated a new GT dinucleotide in intron 41 (Online Supplementary Table S3). The exon 38-43 region showed the expected PCR product in both cell types (772 bp). To confirm these results, two informative SNPs (rs216321 in exon 20 and rs216902 in exon 35) were genotyped. Both were in heterozygous state, indicating that the allele carrying the mutation was unaffected by NMD.

The **c.7730-56C>T** mutation (intron 45) was identified in a type 1 VWD patient (UMP05). The exon 43-49 region, examined by Sanger and NGS, showed no visible effect on mRNA splicing. To confirm the presence of the allele carrying the intronic mutation, two informative SNPs were analyzed: rs216321 and rs1800380. Surprisingly, the sequence obtained in both cell types indicated that only one allele was expressed. To further explore these results, we tested whether intron 45 was retained within the mature mRNA (Online Supplementary Methods) and performed complete sequencing of *VWF* cDNA from leukocytes. Nonetheless, no changes were observed (data not shown).

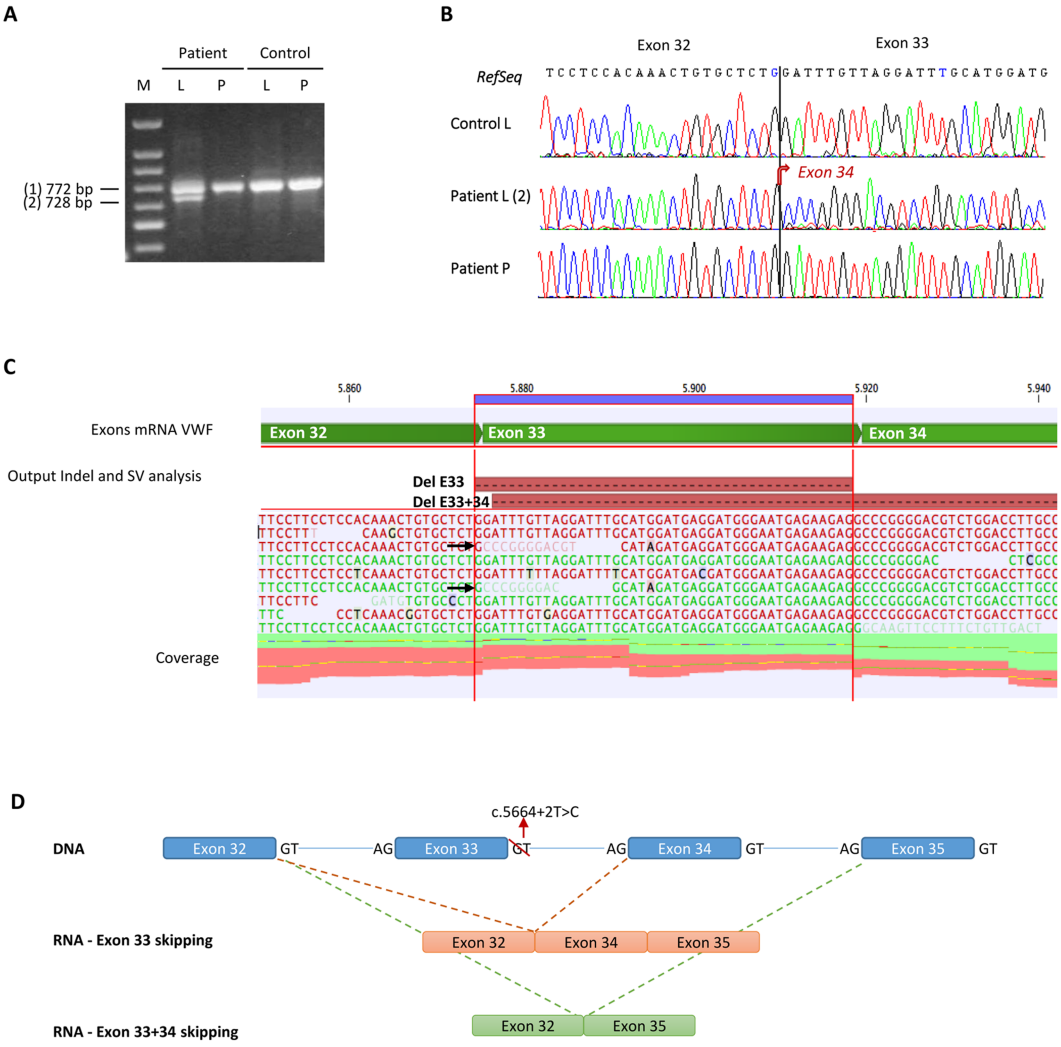


Figure 1. *Analysis of the c.5664+2T>C mutation in patient UMP03. A) Amplified RT-PCR products located in exons 30 and 35 in leukocyte (L) and platelet (P) RNA, separated on 1% agarose gel. B) Traditional Sanger sequencing of PCR products from patient leukocytes (L) and platelets (P), and control leukocytes (L). Analysis of the band 2 from patient leukocyte on agarose gel demonstrates exon 33 skipping. In platelets, only the allele without c.5664+2T>C mutation could be amplified. C) NGS of PCR products from leukocytes showed exon 33 skipping, indicated by arrows depict aberrant transcripts. However, exon 33-34 skipping was also detected, but in a really low of transcripts. D) Schematic representation of the mutation in genomic DNA and its effect on the VWF mRNA sequence. M indicates a 100-bp DNA ladder.*

The **c.7730-4C>G** mutation (intron 45) was identified in a type 2A/2M patient (UMP06), combined with the p.Arg1374Cys mutation. The HSF predicted activation of a cryptic intronic acceptor splice site (ASS), but in the exon 43-49 region, splicing was not affected by the mutation. To confirm the presence of the allele carrying c.7730-4C>G, the p.Arg1374Cys mutation was analyzed, and both nucleotides were seen in platelets and leukocytes, suggesting expression of both alleles. We then performed the same

experiment as was done in patient UMP05 to test intron 45 retention, but no changes in *VWF* cDNA were observed (data not shown).

The **c.8254-5T>G** mutation (intron 51) was found in a type 1 VWD patient (UMP07). The exon 49-52 region was amplified, and the expected 546-bp size was observed in both cell types. Sequence analysis by the 2 techniques revealed no changes in *VWF* cDNA. Interestingly, this patient's 2 children, heterozygous for the c.8254-5T>G mutation, had bleeding symptoms. Hence, we performed complete sequencing of *VWF* cDNA from leukocytes. However, no changes were identified.

Synonymous mutations

The mutations **c.546G>A** (exon 6) and **c.4866C>T** (exon 28) were identified *in trans* in a type 1 VWD patient (UMP08). These mutations were studied only in leukocyte RNA. To investigate c.546G>A, the exon 4-7 region was amplified, which resulted in the expected band and a slightly diffuse, smaller band. Sanger sequencing detected the c.546G>A change (Figure 2). However, NGS not only detected the nucleotide change, but also recognized a loss of 125 nucleotides corresponding to exon 6 (p.Thr179ProfsTer31) in small 2.3% of reads (Online Supplementary Table S4). The predicted impact of the c.546G>A mutation was discrepant and only two *in silico* algorithms predicted an effect on mRNA splicing (Online Supplementary Table S3) that was confirmed *in vivo* even though with a slight effect. To study c.4866C>T, the exon 25-31 region was amplified and sequenced, but no effect on mRNA processing was seen, as was predicted by *in silico* tools. Both nucleotides were identified in heterozygous state, indicating the presence of both alleles. Only offspring carrying the c.546G>A mutation (1 of 3 children) had lower levels of VWF:Ag and VWF:RCo compared to the others. This may be explained by the mutation and/or other factors not explored in the current study.

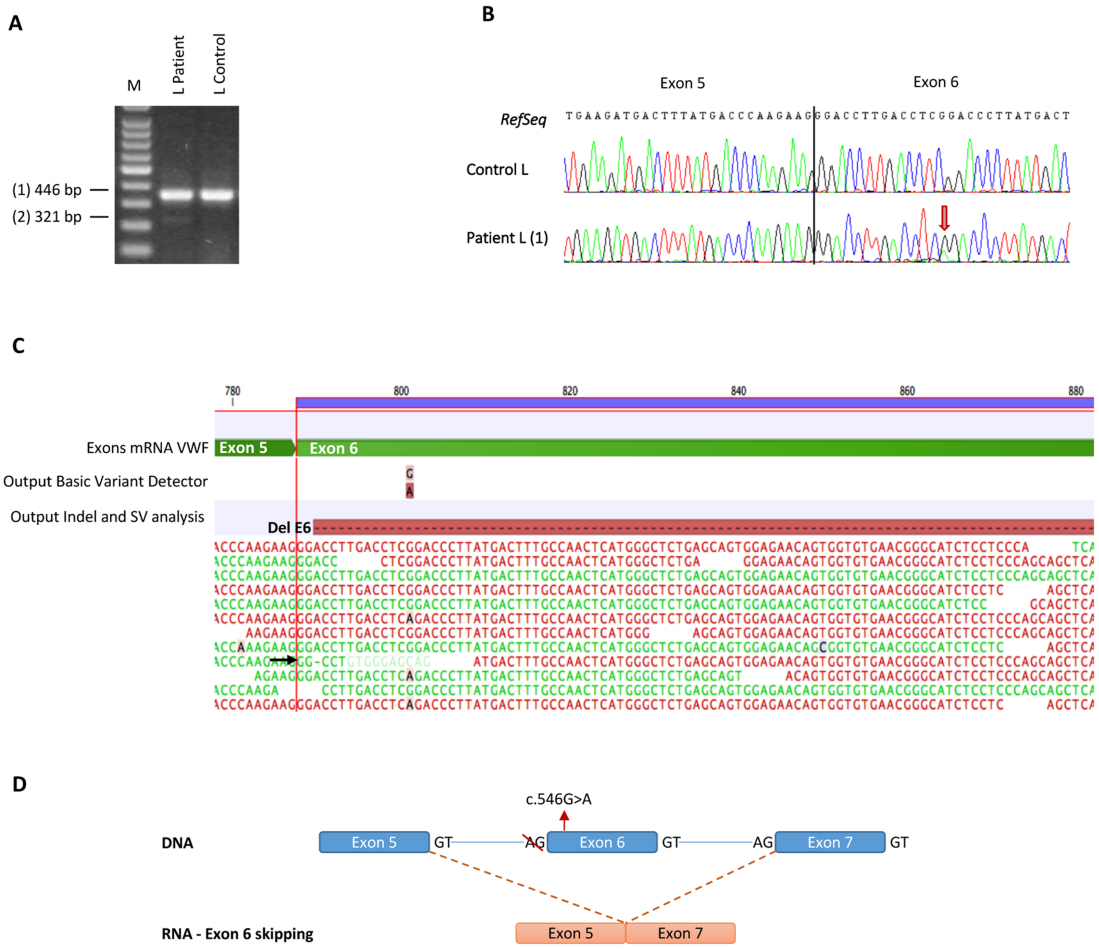


Figure 2. *Analysis of the c.546G>A (p.=) mutation, located at nucleotide 14 from the beginning of exon 6, in patient UMP08. A) Amplified RT-PCR products located in exons 4 and 7 in leukocyte RNA (L), separated on 1% agarose gel. B) Traditional Sanger sequencing of PCR product from patient L (1) showed the single nucleotide change. C) NGS of PCR products from leukocytes identified the single nucleotide variant, as well as skipping of exon 6. Arrows show aberrant transcripts. D) Schematic representation of the mutation in genomic DNA and its effect on the VWF mRNA sequence. M indicates a 100-bp DNA ladder***.

The **c.3291C>T** mutation (exon 25), was identified in patient UMP09 (female), as well as a novel mutation in *F8* (c.1346C>G, p.Ala430Gly). Of note, her brother, who was not included in this study, also had these mutations. As both siblings experienced bleeding, and the HSF prediction regarding c.3291C>T interpreted creation of an exonic splicing silencer site that could potentially alter splicing, we hypothesized that this synonymous mutation could have an effect on *VWF* mRNA processing. Hence, the exon 22-26 region was amplified, but no change in the *VWF* cDNA sequence was observed. In addition, we tested whether intron 25 was retained within mature *VWF* mRNA (Online

Supplementary Methods), but no differences compared to control leukocyte cDNA were found (data not shown). Finally, as both siblings presented similar FVIII:C levels, we investigated X chromosome inactivation in patient UMP09. The results demonstrated skewed inactivation of the WT X chromosome in this patient (Online Supplementary Figure S5).

Missense mutations

The mutation **c.1109G>A** (p.Cys370Tyr, exon 9) was identified in a type 3 VWD carrier (UMP10). The exon 6-12 region was investigated. Whereas the expected size was observed in platelets, amplification from leukocytes resulted in 2 additional bands of ~650 and ~500 bp (Figure 3). Sanger sequencing of leukocyte mRNA confirmed that the 650-bp band corresponded to a transcript lacking exon 9. NGS of leukocyte mRNA revealed 3 different transcripts: the WT, a transcript skipping exon 9 (p.Glu333AlafsTer87) (13% of reads), and a transcript skipping exons 8 and 9 (p.Ser292ThrfsTer87) (7% of reads)(Table 2). Sequencing of PCR products from platelets showed no changes, suggesting that the mutated allele had experienced NMD.

Duplication mutation

The **c.3223-7_3236dup** mutation, which generated an in-frame tandem duplication of 21 nucleotides, including 7 nucleotides from intron 24 and the first 14 from exon 25, was detected in a type 1 VWD patient (UMP11). *S*tudy of this mutation by the 2 techniques demonstrated a change in the native ASS of intron 24 in both cell types, which resulted in maintaining the 21-bp insert corresponding to 7 aberrant amino acids in mature *VWF* mRNA (p.Pro1079_Tyr1080insLeuGlnValAspProGluPro) (Table 2 and Online Supplementary Figure S6). Furthermore, NGS showed that the aberrant mRNA transcript was present in ~14% of reads in both cell types (Online Supplementary Table S4).

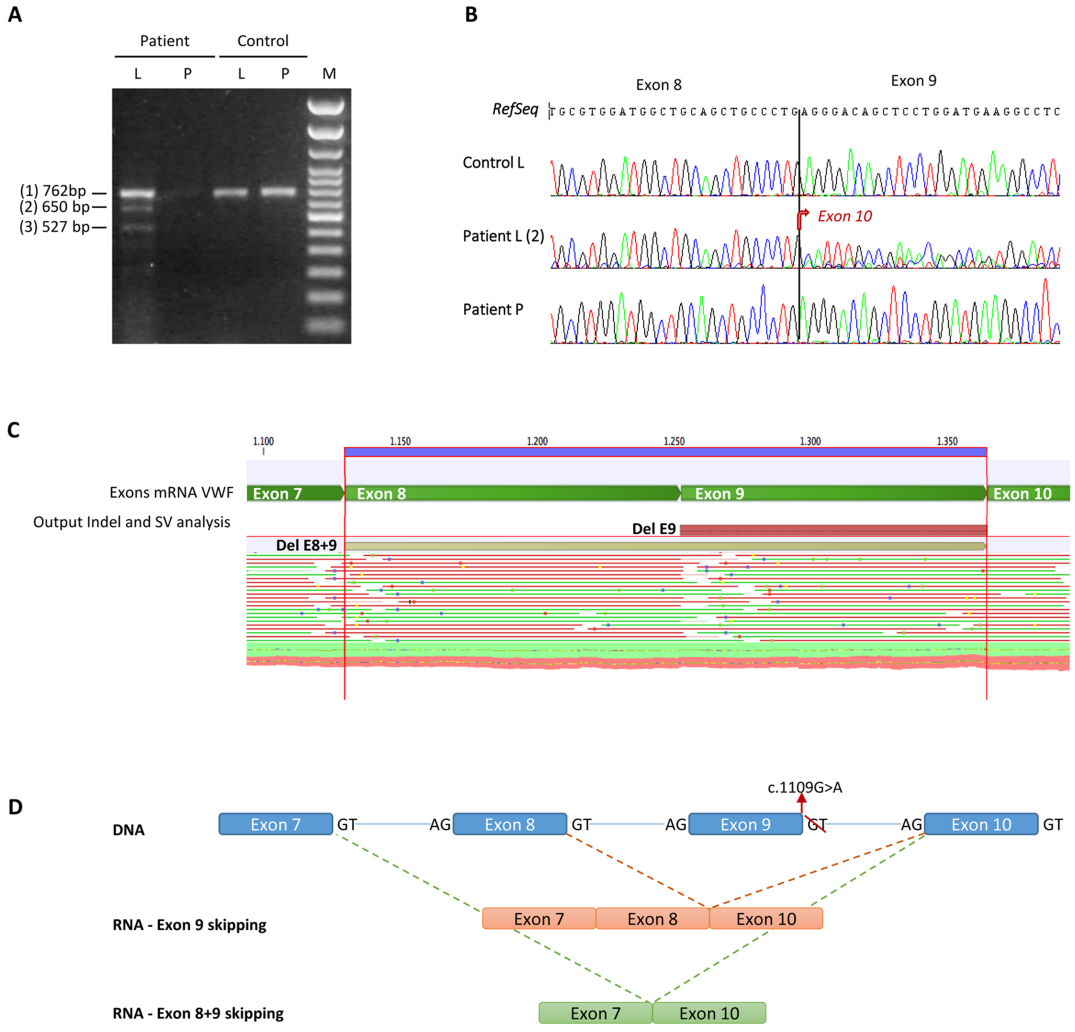


Figure 3. *Analysis of the c.1109G>A (p.Cys370Tyr) mutation in patient UMP10, located in the last nucleotide of exon 9. A) Amplified RT-PCR products located in exons 6 and 12 in leukocyte (L) and platelet (P) RNA, separated on 1% agarose gel. B) Traditional Sanger sequencing of PCR product from patient L (2) showed skipping of exon 9. The L (3) band could not be purified in the agarose gel due to low concentration, thus it was not analyzed by Sanger. In platelets, only the non-mutated allele could be amplified. C) NGS of PCR products from leukocytes identified transcripts lacking exon 8, as well as transcripts lacking exons 8 and 9. D) Schematic representation of the mutation in genomic DNA and its effect on the VWF mRNA sequence. M indicates a 100-bp DNA ladder.*

Combined potential splice site mutations in individual patients

The **c.7082-2A>G** mutation (intron 41) was identified in a type 3 VWD patient (UMP12) combined *in trans* with the **p.Leu150Pro** (c.449C>T, exon 5) mutation. To study c.7082-2A>G, the exon 38-43 region was amplified in platelets and leukocytes, which yielded a band of the expected size. However, Sanger and NGS sequencing of the PCR products revealed activation of a cryptic ASS within exon 42. This led to deletion of the 7 initial nucleotides of exon 42 (r.7082_7088del), as was predicted by the 4 algorithms, and

resulted in a frameshift and a PTC (p.Ala2361GlyfsTer40) (Figure 4). Additionally, NGS showed that the aberrant transcript was present in 24% of reads. On sequencing of platelet amplicons, there were no changes, a finding highly suggestive of NMD. As type 3 VWD is characterized by absent or non-functional expression of both *VWF* alleles, we postulated that the p.Leu150Pro missense mutation could have an effect on *VWF* mRNA. Nonetheless, no sequence change was identified. These results were confirmed by analysis of 2 informative SNPs, rs1800375 and rs1800376, which were found in heterozygous state in leukocyte transcripts and homozygous state in platelets.

The c.[3426T>C; 3485_3486delinsTG] mutations (exon 26) were found in a type 2A VWD patient (UMP13), combined *in trans* with p.Cys2773Ser. *In silico* analysis predicted no impact on the splice site for c. 3485_3486delinsTG, but 2 of the 4 algorithms predicted that c.3426T>C would have an effect on splicing. The exon 25-28 region was investigated, but no effect was found on mRNA processing. Study of the p.Cys2773Ser mutation and 2 informative SNPs (rs2228317 and rs4021576) confirmed that the 2 alleles were present, suggesting that exon 26 mutations do not have an effect on *VWF* splicing.

The mutations **c.7473G>A** (p.=, exon 43), and **c.6699_6702dup** (p.Cys2235Argfs8Ter, exon 38) were identified in *trans* in a severe type 1 VWD patient (UMP14) and studied only in leukocyte RNA. For c.7473G>A, we amplified exons 41-45, and 2 bands emerged: one with the expected size, and a smaller band of~350 bp (Figure 5). Sequencing analysis by both techniques revealed 2 aberrant transcripts: one lacked exon 43 (p.Val2430GlyfsTer335) and the other showed activation of a cryptic ASS within exon 43, leading to deletion of its 4 final nucleotides (p.Ser2479AlafsTer23) (Table 2). Study of c.6699_6702dup in leukocyte mRNA showed no splicing changes, but the duplication introduced 4 nucleotides that changed the reading frame and generated a PTC (p.Cys2235ArgfsTer8).

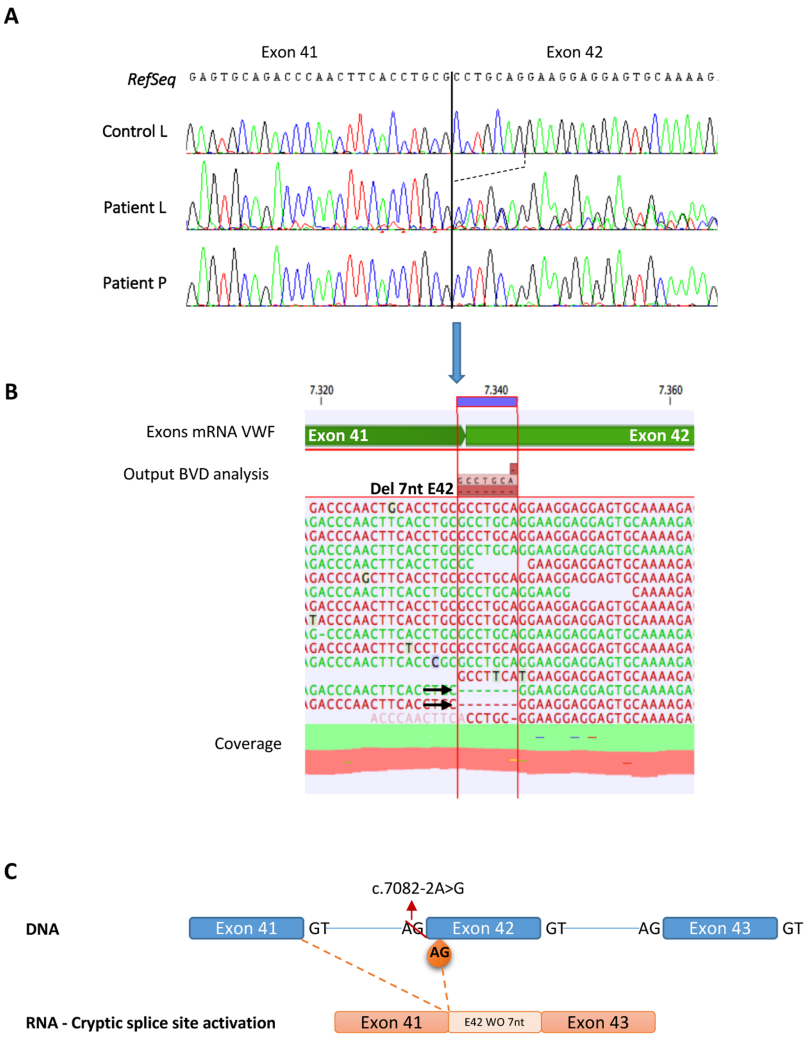


Figure 4. *Analysis of the c.7082-2A>G mutation in patient UMP13. Agarose gel electrophoresis results of RT-PCR amplification of exon 38 to 43 using RNA from leukocytes and platelets were the same as those of healthy controls (data not shown). A) Traditional Sanger sequencing of PCR product from patient leukocyte (L) demonstrate activation of a cryptic splice site 7 nucleotides downstream of the native splice site within exon 42. In patient platelet (P), only the allele carrying the p.Leu150Pro mutation could be amplified. B) NGS of PCR products from leukocytes showed deletion of the 7 initial nucleotides of exon 42. Arrows show aberrant transcripts. C) Schematic representation of the mutation in genomic DNA and its effect on the VWF mRNA sequence. WO indicates without.*

The **c.546G>A** (p.=, exon 6) and **c.7082-2A>G** (intron 41) mutations *in trans* with **c.8155+3G>C** (intron 50) were detected in a type 3 VWD patient (UMP15). This patient had been described in our previous article, but at that time we were only able to identify the effect of c.8155+3G>C (p.Gly2706ValfsTer24) on splicing. On reanalysis of this case with NGS, we were able to characterize the effect of c.546G>A (p.Thr179ProfsTer31) and c.7082-2A>G (p.Ala2361GlyfsTer40) *in cis* in leukocyte

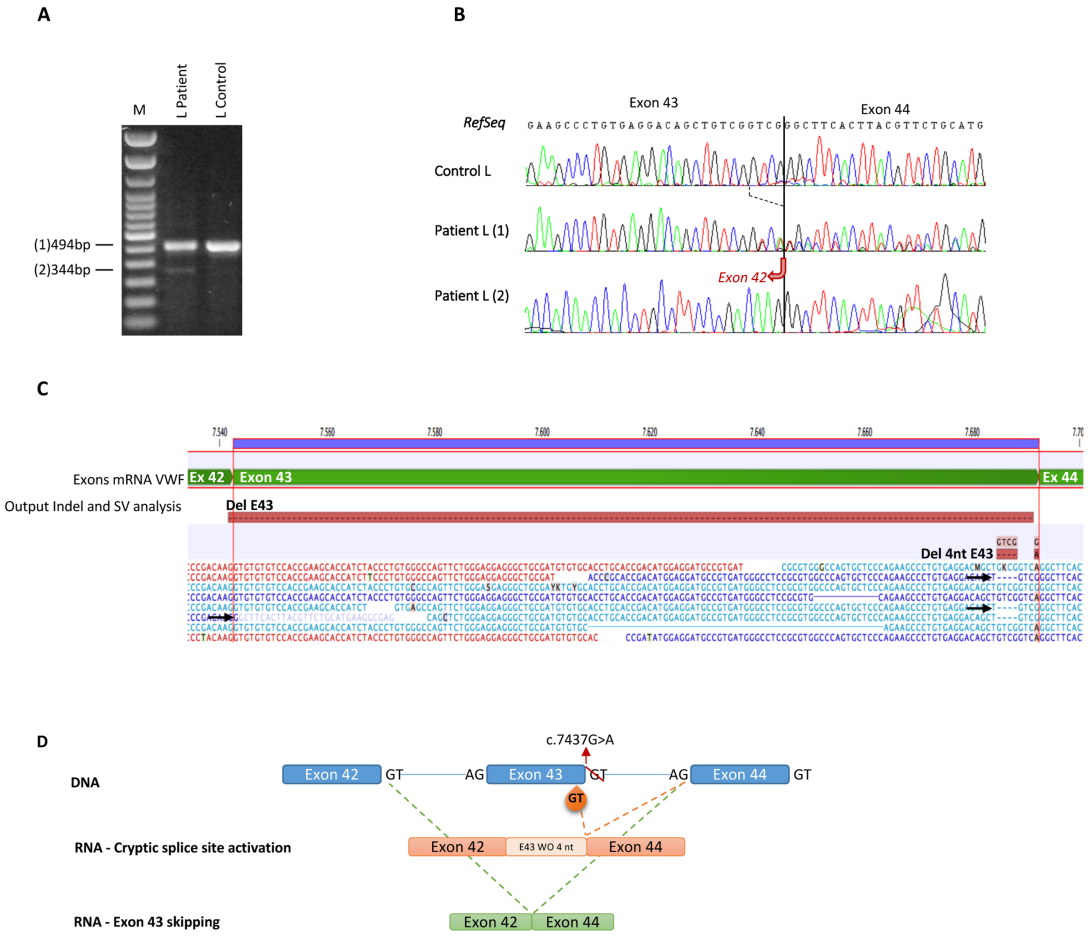


Figure 5. *Analysis of the c.7437G>A mutation in patient UMP15, located in the last nucleotide of exon 43. A) Amplified RT-PCR products located in exons 41 and 45 in leukocyte RNA (L), separated on 1% agarose gel. B) Traditional Sanger sequencing of PCR product from patient leukocyte (L) shows two aberrant transcripts: activation of a cryptic splice site 4 nucleotides upstream within exon 43 (1), and skipping of exon 43 (2). C) NGS of PCR products gave the same results than Sanger sequencing. D) Schematic representation of the mutation in genomic DNA and its effect on the VWF mRNA sequence. M indicates a 100-bp DNA ladder and; WO, without.*

mRNA. The results are consistent with those observed in patients UMP08 and UMP12, who harbored each mutation individually. In addition, in leukocytes, transcripts resulting from c.546G>A were detected in 2% of reads, transcripts from c.7082-2A>G in 6% and transcripts from c.8155+3G>C in 95% of reads. In platelet analyses, NGS detected transcripts without exon 6 resulting from the c.546G>A mutation in 11% of the total reads but no aberrant transcript derived from the mutation c.7082-2A>G indicating that NMD rate could differ between 5' and 3' regions of mRNA. These results indicate the c.[546G>A; 7082-2A>G] allele is under-represented compared to the allele

carrying the c.8155+3G>C mutation, showing that this allele underwent NMD in both cell types, as was previously hypothesized.[7]

DISCUSSION

This study reports the results of in-depth analysis of 18 PSSM in samples from VWD patients. The novel and robust procedure used, combining two-step RT-PCR and NGS sequencing, was faster and more sensitive than the method used in our previous article.[7] This new approach provides several advantages, such as allele-specific individual sequences, higher sensitivity to detect transcript variants present at low copy numbers, and simplified sample preparation. Moreover, the NGS data on total reads obtained for each transcript is additional information that can provide an approximation of the expression levels of each *VWF* mRNA. Studies are currently ongoing to confirm that the relative expression of each transcript obtained by NGS is comparable to that provided by real-time RT-PCR.

The proven usefulness of leukocyte analysis to interpret mutations masked by NMD in platelets was seen in relation to the **c.3379+1G>A** mutation, which had been previously investigated by RT-PCR in platelets.[23] In that study, the true pathogenic effect of this mutation could not be determined due to NMD. In the present study, using leukocyte mRNA and NGS technology, we found that the mutation induces two aberrant transcripts.

Similarly, the pathogenic mechanism of 7 PSSM (c.1533+1G>A, c.5664+2T>C, c.7082-2A>G, c.546G>A, c.7437G>A, and p.Cys370Tyr) was determined in leukocytes, and only the **c.3223-7_3236dup** mutation effect was observed in both cell types. Because this mutation is located in the D3 domain implicated in multimerization, it is possible that the pathogenic mechanism leading to type 1 VWD may be related to impaired VWF secretion due to intracellular retention, as has been described for other D3 mutations such as p.Cys1130Phe.[24] To confirm this hypothesis, expression studies by means of heterologous cell lines or blood outgrowth endothelial cells (BOECs) will remain the gold standard.

Table 2. *Effect of VWF mutations on mRNA from leukocytes and platelets*

Mutation type	NT change	AA change	Exon	Intron	Leukocyte effect	Platelet effect	Protein prediction
Intronic	**c.1533+1G>A**	-	-	13	Exon 13 skipping (r.1433_1533del)	NMD	p.Gly478AlafsTer138
	c.3379+1G>A*	-	-	25	Skipping exon 25 (r.3223_3379del)	NMD	p.Pro1075ValfsTer88
					Activation of a cryptic site 31 nt (r.3349_3379del) ‡	NMD	p.Pro1117ValfsTer88
	c.5664+2T>C	-	-	33	Skipping exon 33 (r.5621_5664del)	NMD	p.Gly1874AlafsTer32
					Exon 33+34 skipping(r.5621_5842del) ‡	NMD	p.Phe1875_Cys1948del
	c.7081+6G>T	-	-	41	*No visible effect*	*No visible effect*	-
	c.7082-2A>G †	-	-	41	Activation of a cryptic site +7 nt in exon 42 (r.7082_7088del)	NMD	p.Ala2361GlyfsTer40
	c.7730-4C>G	-	-	45	*No visible effect*	*No visible effect*	-
	c.7730-56C>T	-	-	45	*No visible effect*	*No visible effect*	-
	c.8254-5T>G	-	-	51	*No visible effect*	*No visible effect*	-
Synonymous	**c.546G>A †**	p.Ser182	6	-	Skipping exon 6 (r.533_657del) ‡	NA	p.Thr179ProfsTer31
	c.3291C>T	p.Cys1097	25	-	*No visible effect*	*No visible effect*	-
	c.3426T>C	p.Cys1142	26	-	*No visible effect*	*No visible effect*	-
	c.4866C>T	p.Asp1622	28	-	*No visible effect*	NA	-
	c.7437G>A	p.Ser2479	43	-	Skipping exon 43 (r.7288_7437)	NA	p.Val2430GlyfsTer335
					Activation of a cryptic site +146 nt in exon 43 (r.7434_7437del)	NA	p.Ser2479AlafsTer23
Missense	**c.449T>C**	p.Leu150Pro	5	-	*No visible effect*	*No visible effect*	-
	c.1109G>A	p.Cys370Tyr	9	-	Skipping exon 9 (r.998_1109del)	NMD	p.Glu333AlafsTer87
					Skipping exon 8+9 (r.875_1109del)	NMD	p.Ser292ThrfsTer87
DelIns	**c.3485_3486delinsTG**	p.Pro1162Leu	26	-	*No visible effect*	*No visible effect*	-
	c.3223-7_3236dup	p.Pro1079_Tyr1080insGluGlnValAspProGluPro	25	-	*(r.3477_3478instttgcaggtggaccccgagcc)*	(r.3477_3478instttgcaggtggaccccgagcc)	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro
	c.6699_6702dup	p.Cys2235ArgTer8	38	-	*No visible effect*	*NMD*	p.Cys2235ArgTer8

AA indicates amino acid; NA, not available; NMD, nonsense-mediated decay and; NT, nucleotide. *Previous functional studies by Nesbitt *et al.*[23] †Previous functional studies by Corrales *et al.*[7] ‡Leukocyte effect observed in a really low % of reads by NG

As would be expected, mutations located in the consensus splicing sequence affected mRNA processing. These included **c.1533+1G>A, c.5664+2T>C,** leading to exon skipping, and **c.7082-2A>G,** activating a cryptic splice site. Both these molecular effects have been reported previously in splicing mutations causing VWD.[7, 25, 26] c.7082-2A>G *in trans* with p.Leu150Pro leads to the development of type 3 VWD. The p.Leu150Pro mutation does not affect splicing. However, other mutations located in the propeptide, such as p.Asp141Tyr, have been identified in type 3 patients,[27] and, as has been described in these cases, we suspect that p.Leu150Pro may compromise propeptide folding and affect intracellular survival and the capacity to mediate multimerization. The c.1533+1G>A mutation in leukocytes generated 3 aberrant transcripts. One of them, which leads to the inframe deletion of exons 13-14, should have been detected in platelets since it would not generate a PTC and would not be degraded. However, because the mutation is located in a region with weak splicing signal sequences, as previously demonstrated,[7, 25, 26] we suggest that the results obtained in leukocytes could be an artifact that would not occur in the natural cellular type of VWF expression. Based on these observations, we propose that the real effect of c.1533+1G>A is the skipping of exon 13, which predicts a frameshift at position 478 and addition of 138 aberrant amino acids before a PTC is encountered (p.Gly478AlafsTer138).

Two synonymous mutations included in our study, **c.546G>A** and **c.7437G>A**, affect splicing of *VWF* mRNA. To our knowledge, only 3 synonymous mutations affecting *VWF* mRNA processing have been reported: c.7056C>T,[28] c.7464C>T,[29] and c.3390C>T.[26] Of note, the results of the present study have almost doubled the number of reported mutations of this type causing VWD. The pathogenic effect of **c.546G>A** could only be documented by NGS; it eluded Sanger detection because of the low expression in leukocytes and platelets. The **c.7437G>A** mutation located at the last nucleotide of exon 43 was found in a patient with severe type 1 VWD (UMP15) *in trans* with c.6699_6702dup (p.Cys2235ArgTer8). This synonymous mutation produced 2 splicing variants. First, creation of a new DSS 4 nucleotides upstream of the WT DSS in exon 43 that generated a PTC in exon 44 and would lead to NMD in platelets; and second, skipping of exon 43 and generation of a PTC in exon 52. In this latter case, we would

expect that NMD had been abolished in platelets, since this cellular mechanism is not effective when a PTC is encountered within 50 bp of the last exon-exon junction.[30]

Study of the **p.Cys370Tyr** (c.1109G>A in exon 9) mutation, resulted in generation of 2 *VWF* transcripts, one of them identified only by NGS because of their low expression. Of note, only one other missense mutation, p.Gly1108Arg, has been reported to affect *VWF* splicing. Thus, these 2 mutations support the concept that missense mutations on the last exonic nucleotides can also have repercussions on the splicing process.[31] Interestingly, the pathogenic effect of this mutation is the same as that described for c.1109+2T>C (intron 9).[32] In addition to the patient reported here, the p.Cys370Tyr mutation was identified in 6 additional related patients included in the PCM-EVW-ES: in homozygous state in a type 3 VWD patient and in heterozygosis in 4 type 3 carriers and 1 patient with type 1 VWD, based on their phenotype levels. Certain genetic modifiers of VWF levels[33] and inter-individual variability in NMD efficiency between patients carrying identical mutations may lead to differences in the disease severity and clinical phenotype.[34]

Lastly, study of c.7081+6G>T, c.7730-4C>G, c.7730-56C>T, c.8254-5T>G, c.3291C>T, c.4866C>T and c.[3426T>C; 3485_3486delinsTG] showed no visible effect on mRNA processing, although this does not necessarily mean that they have no effect on splicing. For instance, mRNA SNP analysis in patient UMP05 with **c.7730-56C>T** showed an absence of 1 allele in leukocytes and platelets, suggesting that the allele may have experienced NMD in both cell types or lack expression due to a mutation that was not detected by our sequencing protocol. Of particular note, deep intronic mutations such as c.6599-20A>T have been found to cause VWD.[35] In patient UMP07 carrying the **c.8254-5T>G** mutation**,** no informative SNP was identified and we were unable to determine whether there was a lack of expression of 1 allele, which could explain cosegregation of the mutation in the family with bleeding symptoms.

Study of **c.[3426T>C; 3485_3486delinsTG]** and **c.7730-4C>G** showed no visible effect on splicing. However, these mutations have been identified in homozygous state in a type 3 patient (c.[3426T>C; 3485_3486delinsTG])[3] and in a severe type 1 VWD patient

(c.7730-4C>G)[4] with VWF:Ag at 7% and VWF:RCo at 5%, combined with the heterozygous p.Ala631Val (previously reported in a healthy control).[36] Based on these findings, there is substantial evidence that these mutations would have an effect on VWF levels. Therefore, to unequivocally determine the potential deleterious effect of these variants, functional studies remain essential. These studies are traditionally undergoing by *in vitro* analyses performed using heterologous cell lines (COS7, AtT-20 and HEK293). However, the advent of the possibility of obtaining BOECs from patients represents a valuable alternative, since this is the functional expression site of *VWF*.[37] Moreover, BOECs allows protein expression and mRNA studies simultaneously.

In silico algorithms used to assess the impact of mutations on splicing are more sensitive and accurate in determining the putative effect of intronic mutations than that of synonymous or missense mutations, such as c.546G>A and p.Cys370Tyr. Therefore, the results here should not be considered definitive, and as with all analytical approaches, should form one aspect of a wider investigation.[5]

In conclusion, we present an extensive study reporting the effect of 18 candidate mutations on *VWF* mRNA processing. *In vivo* mRNA studies incorporating NGS technology together with traditional sequencing enabled us to determine the pathogenic effect of 8 PSSM (44%). Our study emphasizes the importance of examining selected mutations, including synonymous and missense mutations, to determine their pathogenic role in splicing. Taken together, our results add to the current knowledge about the molecular events leading to VWD.

Acknowledgments

We are indebted to Baxalta US Inc., now a part of Shire, for support of the PCM-EVW-ES (Grant H13-000845). This study was also supported by the Spanish Ministry of the Economy and Competitiveness (MINECO, *Ministerio de Economía y Competitividad)*, Instituto de Salud Carlos III (ISCIII) (PI12/01494, PI15/01643 and RD12/0042/0053). We are very grateful for the kind collaboration of the participating patients and their

families. CIBERCV is an initiative of ISCIII, co-financed by the European Regional Development Fund (ERDF), “A way to build Europe”.

REFERENCES

1. Rodeghiero F, Castaman G, Dini E. Epidemiological investigation of the prevalence of von Willebrand's disease. Blood. 1987;69(2):454-459.
2. Batlle J, Perez-Rodriguez A, Corrales I, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): Proposal for a new diagnostic paradigm. Thromb Haemost. 2016;115(1):40-50.
3. Borras N, Batlle J, Perez-Rodriguez A, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): comprehensive genetic analysis by next-generation sequencing of 480 patients. Haematologica. 2017;102(12):2005-2014.
4. Fidalgo T, Salvado R, Corrales I, et al. Genotype-phenotype correlation in a cohort of Portuguese patients comprising the entire spectrum of VWD types: impact of NGS. Thromb Haemost. 2016;116(1):
5. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-424.
6. Wang QY, Song J, Gibbs RA, Boerwinkle E, Dong JF, Yu FL. Characterizing polymorphisms and allelic diversity of von Willebrand factor gene in the 1000 Genomes. J Thromb Haemost. 2013;11(2):261-269.
7. Corrales I, Ramirez L, Altisent C, Parra R, Vidal F. The study of the effect of splicing mutations in von Willebrand factor using RNA isolated from patients' platelets and leukocytes. J Thromb Haemost. 2011;9(4):679-688.
8. Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. Nature reviews. 2007;8(10):749-761.
9. Eikenboom JC, Reitsma PH, Peerlinck KM, Briët E. Recessive inheritance of von Willebrand's disease type I [see comments]. Lancet. 1993;341(8851):982-986.
10. Jaffe EA, Hoyer LW, Nachman RL. Synthesis of von Willebrand factor by cultured human endothelial cells. Proc Natl Acad Sci U S A. 1974;71(5):1906-1909.
11. Sporn LA, Chavin SI, Marder VJ, Wagner DD. Biosynthesis of von Willebrand protein by human megakaryocytes. J Clin Invest. 1985;76(3):1102-1106.
12. Kieffer N, Guichard J, Farcet JP, Vainchenker W, Breton-Gorius J. Biosynthesis of major platelet proteins in human blood platelets. Eur J Biochem. 1987;164(1):189-195.

13. Fink L, Holschermann H, Kwapiszewska G, et al. Characterization of platelet-specific mRNA by real-time PCR after laser-assisted microdissection. Thromb Haemost. 2003;90(4):749-756.

14. McVey JH, Boswell EJ, Takamiya O, et al. Exclusion of the first EGF domain of factor VII by a splice site mutation causes lethal factor VII deficiency. Blood. 1998;92(3):920-926.

15. David D, Santos IM, Johnson K, Tuddenham EG, McVey JH. Analysis of the consequences of premature termination codons within factor VIII coding sequences. J Thromb Haemost. 2003;1(1):139-146.

16. Roberts RG, Bentley DR, Bobrow M. Infidelity in the structure of ectopic transcripts: a novel exon in lymphocyte dystrophin transcripts. Hum Mutat. 1993;2(4):293-299.

17. Byers PH. Killing the messenger: new insights into nonsense-mediated mRNA decay. J Clin Invest. 2002;109(1):3-6.

18. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci U S A. 2003;100(1):189-192.

19. Martorell L, Luce E, Vazquez JL, et al. Advanced cell-based modeling of the royal disease: characterization of the mutated F9 mRNA. J Thromb Haemost. 2017;15(11):2188-2197.

20. Hojny J, Zemankova P, Lhota F, et al. Multiplex PCR and NGS-based identification of mRNA splicing variants: Analysis of BRCA1 splicing pattern as a model. Gene. 2017;637(41-49.

21. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. J Mol Biol. 1991;220(1):49-65.

22. Corrales I, Ramirez L, Altisent C, Parra R, Vidal F. Rapid molecular diagnosis of von Willebrand disease by direct sequencing. Detection of 12 novel putative mutations in VWF gene. Thromb Haemost. 2009;101(3):570-576.

23. Nesbitt IM, Hampton KK, Preston FE, Peake IR, Goodeve AC. A common splice site mutation is shared by two families with different type 2N von Willebrand disease mutations. Thromb Haemost. 1999;82(3):1061-1064.

24. Tjernberg P, Vos HL, Castaman G, Bertina RM, Eikenboom JC. Dimerization and multimerization defects of von Willebrand factor due to mutated cysteine residues. J Thromb Haemost. 2004;2(2):257-265.

25. Gallinaro L, Sartorello F, Pontara E, et al. Combined partial exon skipping and cryptic splice site activation as a new molecular mechanism for recessive type 1 von Willebrand disease. Thromb Haemost. 2006;96(6):711-716.

26. Pagliari MT, Baronciani L, Garcia Oya I, et al. A synonymous (c.3390C>T) or a splice-site (c.3380-2A>G) mutation causes exon 26 skipping in four patients with von Willebrand disease (2A/IIE). J Thromb Haemost. 2013;11(7):1251-1259.

27. Baronciani L, Federici AB, Cozzi G, et al. Expression studies of missense mutations p.D141Y, p.C275S located in the propeptide of von Willebrand factor in patients with type 3 von Willebrand disease. Haemophilia. 2008;14(3):549-555.

28. Daidone V, Gallinaro L, Grazia Cattini M, et al. An apparently silent nucleotide substitution (c.7056C>T) in the von Willebrand factor gene is responsible for type 1 von Willebrand disease. Haematologica. 2011;96(6):881-887.

29. Yadegari H, Biswas A, Akhter MS, et al. Intron retention resulting from a silent mutation in the VWF gene that structurally influences the 5' splice site. Blood. 2016;128(17):2144-2152.

30. Maquat LE. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nature reviews Molecular cell biology. 2004;5(2):89-99.

31. James PD, O'Brien LA, Hegadorn CA, et al. A novel type 2A von Willebrand factor mutation located at the last nucleotide of exon 26 (3538G>A) causes skipping of 2 nonadjacent exons. Blood. 2004;104(9):2739-2745.

32. Castaman G, Plate M, Giacomelli SH, Rodeghiero F, Duga S. Alterations of mRNA processing and stability as a pathogenic mechanism in von Willebrand factor quantitative deficiencies. J Thromb Haemost. 2010;8(12):2736-2742.

33. James PD, Lillicrap D. von Willebrand disease: clinical and laboratory lessons learned from the large von Willebrand disease studies. Am J Hematol. 2012;87 Suppl 1(S4-11.

34. Miller JN, Pearce DA. Nonsense-mediated decay in genetic disease: friend or foe? Mutation research Reviews in mutation research. 2014;762(52-64.

35. Hawke L, Bowman ML, Poon MC, Scully MF, Rivard GE, James PD. Characterization of aberrant splicing of von Willebrand factor in von Willebrand disease: an underrecognized mechanism. Blood. 2016;128(4):584-593.

36. Bellissimo DB, Christopherson PA, Flood VH, et al. VWF mutations and new sequence variations identified in healthy controls are more frequent in the African-American population. Blood. 2012;119(9):2135-2140.

37. Wang JW, Bouwens EA, Pintao MC, et al. Analysis of the storage and secretion of von Willebrand factor in blood outgrowth endothelial cells derived from patients with von Willebrand disease. Blood. 2013;121(14):2762-2772.

Material Suplementari Article 4

SUPPLEMENTAL METHODS

VWF mRNA amplification

The primers used have been reported previously,[1] except for those specifically designed to analyze certain mutations (Online Supplementary Table S1). The PCR mix contained 1X PCR Buffer, 1.5 mM $MgCl_2$, 200 µM dNTPs, 2 U Platinum Taq DNA polymerase, 0.75 µM of each primer, and cDNA in a total volume of 25 µL. After initial denaturation at 94ºC for 3 min, 38 cycles of 94ºC for 20 sec, 62ºC for 30 sec, and 72ºC for 1 min were performed, followed by a final extension at 72ºC for 3 min.

CLC Bio Data analysis parameters

Primer sequences and low-quality regions were trimmed and aligned against the WT *VWF* mRNA sequence with specific parameter settings. The Basic Variant Detection algorithm was applied for single nucleotide variant-calling and the Indels and Structural Variants tool was performed to identify insertions and deletions produced by alternative splicing due to mutations. Analytical settings of both algorithms are described in Online Supplementary Figure S1. Finally, a specific probe was designed to each pre-characterized aberrant transcript and used to perform a second alignment from the original dataset. The results indicate the number of reads that aligned to each mRNA probe.

Intron retention of exon 25

To determine whether intron 25 was retained within the mature mRNA of patient UMP09, RT-PCR reactions using 2 allele-specific primer combinations were performed. In the first primer pair, the forward was designed at the junction of exon 24-25 (VWFR24/25-1) and the reverse primer was designed inside intron 25 (VWF_E25-2). In the second primer pair, the forward primer targeted intron 25 (VWF_E26-1) and the reverse was located across the junction of exon 26-27 (VWFR26/27-2).

Intron retention of exon 45

To determine whether intron 45 was retained within the mature mRNA of patients UMP05 and UMP06, RT-PCR reactions using 2 allele-specific primers combinations were performed. In the first primer pair, the forward was designed at the junction of exon 44-45 (VWFR44/45-1C) and the reverse primer was designed inside intron 45 (VWF_E45-2). In the second primer pair, the forward primer targeted intron 45 (VWF_E46-1) and the reverse was located across the junction of exon 47-48 (VWFR46/47-2).

X-chromosome inactivation

The X-chromosome inactivation profile studies were performed on DNA extracted from peripheral blood. The study was done using HpaII digestion. Genomic DNA samples were digested with the methylation-sensitive enzyme HpaII (Invitrogen) and subjected to PCR amplification of the highly polymorphic CAG repeat sequence in the first exon of the human androgen receptor gene (HUMARA), with specific fluorescent primers. The primer sequences are shown in Online Supplementary Table S1. The PCR products, both before and after HpaII digestion, were electrophoresed on an automated DNA sequencer (ABI 3130XL; Applied Biosystems, Warrington, United Kingdom) and analyzed by GeneMapper software (Applied Biosystems).[2]

SUPPLEMENTAL TABLES

Table S1. Specific primers used for amplification and sequencing by Sanger of VWF RNA.

Forward Primer	Location	Reverse Primer	Location	Purpose	Product size
VWRNA1-1 CTACATAACAGCAAGACAGTCC	Exon 1	**VWRNA1-2** GGTGCTCTTCAGAAGCTGG	Exon 7	Full-length amplification VWF	836 bp
VWRNA2-1 GCAGCTCATGCAACATCTCC	Exon 6	**VWRNA2-2** CCTTGTGAAACTGAAGCATGG	Exon 12	Full-length amplification VWF	762 bp
VWRNA3-1 CCATTGTCATTGAGACTGTCCAC	Exon 11	**VWRNA3-2** CCTGACCTGCCGCTCTCT	Exon 16	Full-length amplification VWF	744 bp
VWRNA4-1 GCGTGGCGCGAGCCAGG	Exon 15	**VWRNA4-2** GTAACCCTGGGACCTTTCG	Exon 21	Full-length amplification VWF	796 bp
VWRNA5-1 CATGGCCCACTACCTCACC	Exon 20	**VWRNA5-2** GGTGGTGACCTGGAGGAC	Exon 25	Full-length amplification VWF	752 bp
VWRNA6-1 TCCTGTGAGTCCATTGGGG	Exon 25	**VWRNA6-2** GGAAGCGACCGTCAGAGC	Exon 28	Full-length amplification VWF	750 bp
VWRNA7-1 CTTTGTGGTGGACATGATGG	Exon 28	**VWRNA7-2** CCTCTCTGACCACAGCTTC	Exon 28	Full-length amplification VWF	876 bp
VWRNA8-1 CCTGCAGCGGGTGCGAG	Exon 28	**VWRNA8-2** CCTGGTCACGGACGTCTC	Exon 31	Full-length amplification VWF	728 bp
VWRNA9-1 AAATCGGGGATGCCTTGGG	Exon 30	**VWRNA9-2** GGAGGTGACGGTGAATGGG	Exon 35	Full-length amplification VWF	772 bp
VWRNA10-1 GGAGGTGATTCTCCATAATGG	Exon 35	**VWRNA10-2** CTGCGTTGTGCAGTTGACC	Exon 39	Full-length amplification VWF	930 bp
VWRNA11-1 GTGAGCATGGCTGTCCCC	Exon 38	**VWRNA11-2** GGATGCCGTGATGGGCCT	Exon 43	Full-length amplification VWF	753 bp
VWRNA12-1 GTGTGTCCACCGAAGCACC	Exon 43	**VWRNA12-2** CCCTTTGATGAACACAAGTGT	Exon 49	Full-length amplification VWF	813 bp
VWRNA13-1 CGTGATGAGACGCTCCAGG	Exon 49	**VWRNA13-2** CCTCTGCATGTTCTGCTCTT	Exon 52	Full-length amplification VWF	546 bp
VWRNA3-1 CCATTGTCATTGAGACTGTCCA	Exon 11	**VWRNA_E15-2** CGTCGTAGCGGCAGTTCC	Exon 15	primers to avoid alternative splicing in leukocytes	564 bp
VWFR4-1 GCCTCTCCGTGTATCTTGG	Exon 4	**VWRNA1-2B** GGTGCTCTTCAGAAGCTGG	Exon 7	Exon 5-6	446 bp
VWF22-1 GGTCCTGAAGCAGACATACC	Exon 22	**VWF26-2A** CTCCCGGAGATTCCTCTCC	Exon 26	Exon 25 complete	466 bp
VWFR24/25-1 GACTGCAACAAGCTGGTGG	Junction exon 24-25	**VWF_E25-2** ATCCAGTCCCTACTAACACT	Intron 25	Intron 25 retention	301 bp
VWF_E26-1 ACATAGCAAGACCCCATCTG	Intron 25	**VWFR26/27-2** CATCCAGGATTTTCCCTGG	Junction exon 26-27	Intron 25 retention	401 bp
VWFR43/44-1 GGACAGCTGTCGGTCGGG	Junction exon 43-44	**VWF_E45-2** CAGGAGCCAAAAGTGGAAAG	Intron 45	Intron 45 retention	381 bp
VWF_E46-1 CGACCGATACAGGAGGGAG	Intron 45	**VWF47/48-2** ATTTTCTTCCTTGTAACCCAGG	Junction exon 47-48	Intron 45 retention	295 bp
VWFR41/42-1 CCCAACTTCACCTGCGCC	Junction exon 41-42	**VWFR44/45-2** TGGGAGCCGACACTCTTCC	Junction exon 44-45	Exon 43 complete	494 bp

At the top of the table, primers previously designed by Corrales et al.[1] At the bottom of the table, primers newly designed to analyze potential splice site mutations included in this study

Table S2. Additional laboratory and clinical data of VWD patients.

Patients code	VWD	Hospital	Registry code	Age	Sex	ABO	Multimeric pattern	BS	Family history
UMP01	3 carrier	CHUC	P64's mother	37	F	A	ND	0	Yes
UMP02	1H	CHUAC	C01P034F17	32	M	O+	NORMAL	2	Yes
UMP03	1	CHUAC	C01P080	40	M	O+	ND	1	NA
UMP04	1	HUVH	-	38	F	O+	ND	C, D	No
UMP05	1	CHUAC	C01P006F03	39	F	O+	NORMAL	11	No
UMP06	2A/2M	CHUAC	C01P024F12	33	F	O	SMEAR	10	No
UMP07	1	HUVH	-	47	M	O+	ND	B	Yes
UMP08	1	HUVH	-	42	M	O+	ND	B	Yes
UMP09	1	HUVH	-	20	F	ND	ND	B, C	Yes
UMP10	3 carrier	HUVH	-	45	F	O	ND	A, C, E	Yes
UMP11	1	CHUAC	C01P051F11	36	M	O+	NORMAL	7	Yes
UMP12	3	HUVH	C03P021F25	15	M	A	ND	10	Yes
UMP13	2A	HUVH	C03P048F231	15	F	O+	↓HMWM	0	Yes
UMP14	1	CHUC	P12	10	M	ND	ND	16	Yes
UMP15	3	HUVH	-	25	F	O	ND	B, E	Yes

Bleeding scores calculated with the International Society on Thrombosis and Haemostasis bleeding assessment tools (ISTH-BAT) were used to assess bleeding in patients included in the Spanish (PCM-EVW-ES) and Portuguese registries. In the remaining patients, symptoms were assigned letters: A, easy bruising; B, epistaxis; C, prolonged bleeding from wounds; D, menorrhagia; and E, postoperative bleeding. BS indicates bleeding score; CHUAC, Complexo Hospitalario Universitario A Coruña, Spain; CHUC, Centro Hospitalar e Universitário de Coimbra, Portugal; F, female; HUVH, Vall d'Hebron University Hospital, Barcelona, Spain; HMWM, high molecular weight multimers; M, male; NA; not available; ND, not determined and; VWD, von Willebrand disease.

Table S3. Summary of the *in silico* analysis of PSSM

Mutation Type	NT change	AA change	Exon	Intron	Affected splice site	Adjacent nucleotides	NetGene2		Neural Network Splice		MaxEnt		HSF		Score	HSF Interpretation
							wt	var	wt	var	wt	var	wt	var		
Intronic	c.1533+1G>A	-	-	13	DSS intron 13	GAAGgtaggt>GAAGataggt	1	Native DSS destroyed	0.98	Native DSS destroyed	10.29	2.1	90.71	63.87	4/4	Alteration of the WT donor site, most probably affecting splicing.
	c.3379+1G>A	-	-	25	DSS intron 25	TGCCgtgag>TGCCatgag	0.95	Native DSS destroyed	0.90	Native DSS destroyed	8.56	0.38	81.32	54.48	4/4	Alteration of the WT donor site, most probably affecting splicing.
	c.5664+2T>C	-	-	33	DSS intron 33	AGGgtaagt>AGGgcaagt	0.83	Native DSS destroyed	1	Native DSS destroyed	10.45	2.7	94.19	67.35	4/4	Alteration of the WT donor site, most probably affecting splicing.
	c.7081+6G>T	-	-	41	DSS intron 41	taaggcctc>taagtcctc	1	1	1	1	9.63	11.37	91.56	93.49	0/4	No significant splicing motif alteration detected.
	c.7082-2A>G	-	-	41	ASS intron 41	cctcagCCT>cctcggCCT	NP	Native ASS destroyed	0.69	Native ASS destroyed	6.68	-1.27	88.28	59.34	4/4	Alteration of the WT donor site, most probably affecting splicing.
	c.7730-4C>G	-	-	45	ASS intron 45	tcccccagA>tcccgcagA	1	1	0.99	0.99	11.71	13.00	91.73	91.11	0/4	Activation of an intronic cryptic acceptor site. Potential alteration of splicing.
	c.7730-56C>T	-	-	45	ASS intron 45	tcgccggtc>tcgctggtc	1	1	0.99	0.99	6.54	7.02	87.43	88.69	0/4	Creation of an intronic ESE site. Probably no impact on splicing.
	c.8155+3G>C	-	-	50	DSS intron 50	ATgtgagtg>ATgtcagtg	0.54	Native DSS destroyed	0.94	Native DSS destroyed	7.83	2.02	85.46	78.65	3/4	Alteration of the WT donor site, most probably affecting splicing.
	c.8254-5T>G	-	-	51	ASS intron 51	ttcttgcag>ttctggcag	1	0.96	0.99	0.94	12.47	10.09	92.5	88.8	0/4	No significant splicing motif alteration detected.
Synonymous	c.546G>A	p.Ser182	6	-	New ASS	CCTCGGACC>CCTCAGACC	0.71	0.33	0.90	0.88	NP	NP	57.31	86.26	2/4	Activation of an exonic cryptic acceptor site, with presence of one or more cryptic branch point(s). Potential alteration of splicing.
	c.3291C>T	p.Cys1097	25	-	New DSS	ACTGCGCCT>ACTGTGCCT	0.95	0.95	0.90	0.90	NP	NP	NP	NP	0/4	Creation of an exonic ESS site. Potential alteration of splicing.
	c.3426T>C	p.Cys1142	26	-	Alteration GT	AGTGTGAGT>AGTGCGAGT	0.67	0.67	1	1	6.96	-0.79	82.16	55.33	2/4	Alteration of an exonic ESE site. Potential alteration of splicing.
	c.4866C>T	p.Asp1622	28	-	Not affected	GAGACATCC>GAGATATCC	0.96	0.96	0.97	0.97	NP	NP	75.04	75.88	0/4	Alteration of an exonic ESE site.Potential alteration of splicing.
	c.7437G>A	p.Ser2479	43	-	DSS intron 43	GGTCGgtgag>GGTCAgtgag	1	0.95	0.98	0.63	3.52	11.11	92.11	81.53	3/4	Alteration of the WT donor site, most probably affecting splicing.
Missense	c.449T>C	p.Leu150Pro	5	-	Not affected	CTGCTGTCA>CTGCCGTCA	1	1	0.98	0.98	NP	NP	85.99	86,64	0/4	No significant splicing motif alteration detected.
	c.1109G>A	p.Cys370Tyr	9	-	DSS intron 9	ACCTGgtaa>ACCTAgtaa	NP	NP	0.31	Native DSS destroyed	5.67	-2.74	81.00	70.42	3/4	Alteration of the WT donor site, most probably affecting splicing.
DelIns	c.3485_3486delinsTG	p.Pro1162Leu	26	-	Not affected	GAGCCACTGG>GAGCTGCTGG	0.67	0.83	1	1	6.69	7.09	71.11	71.95	1/4	Creation of an exonic ESS site. Potential alteration of splicing.
	c.3223-7_3236dup	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	25	-	New ASS	cagG-------TGGA>cagGtttgcagGTGGA	0.95	0,95 + New ASS (0.81)	0.73	0,59 + New ASS	3.09	4.86	15.42	84.46	4/4	Activation of an exonic cryptic acceptor site. Potential alteration of splicing.
	c.6699_6702dup	p.Cys2235Argfs*8	38	-	Not affected	AGGC----TGTT>AGGCAGGCTGTT	0,93	0,93	0,97	0,98	4,3	4,27	82,59	83,23	0/4	Alteration of an exonic ESE site.Potential alteration of splicing.

The *in silico* global score is based on the number of *in silico* algorithms that predicted a splicing effect. *In silico* analysis of mutations included in the PCM-ES-EVW registry was performed by ALAMUT. *In silico* analysis of the mutations not included in the PCM-EVW-ES registry was performed individually under the established parameters of each algorithm, as well as NetGene2 (unavailable in ALAMUT). For HSF and MaxEnt, the sequence analyzed is 50 nucleotides in length. For NetGene2 and Neural Network Splice, the sequence analyzed is the exon length plus 100 intronic nucleotides at the exon ends. In the HSF software, if the WT score is above the threshold (65) and the score variation (between WT and Mutant) is below -10%, the mutation is considered to break the splice site. In the other case, if the WT score is below the threshold and the score variation is above +10% the mutation is considered to create a new splice site. In the MaxEnt software, if the WT score is above the threshold (3) and the score variation (between WT and Mutant) is below -30% the mutation is considered to break the splice site. In the other case, if the WT score is below the threshold and the score variation is above +30% the mutation is considered to create a new splice site. AA indicates amino acid; ASS, acceptor splice site; DSS, donor splice site; ESS, exonic splicing silencer; ESE, exonic splicing enhancer; HSF, Human Splicing Finder; MaxEnt, Maximum Entropy; NP not predicted; NT, nucleotide and; PSSM, potential splice site mutation.

Table S4. Analysis of mRNA aberrant transcript expression by next-generation sequencing.

Patient code	Mutation	mRNA probes	Leukocytes		Platelets	
			% Patient Reads	% Control Reads	% Patient Reads	% Control Reads
UMP01	c.1533+1G>A	WT	**28 %**	100 %	**100 %**	100 %
		Exon 13 skipping (r.1433_1533del)	**19 %**	0 %	**0 %**	0 %
		Exon 13+14 skipping (r.1433_1729del)	**45 %**	0 %	**0 %**	0 %
		Exon 14 skipping (r.1534_1729del)	**8 %**	0 %	**0 %**	0 %
UMP02	c.3379+1G>A	WT	**56 %**	100 %	**99 %**	100 %
		Exon 25 Skipping (r.3223_3379del)	**43 %**	0 %	**1 %**	0 %
		Activation of a cryptic site +126 nt in exon 25 (r.3349_3379del) ‡	**1.4 %**	0 %	**0 %**	0 %
UMP03	c.5664+2T>C	WT	**61 %**	100 %	**97 %**	100 %
		Exon 33 skipping (r.5621_5664del)	**37 %**	0 %	**0 %**	0 %
		Exon 33+34 skipping(r.5621_5842del) ‡	**2 %**	0 %	**3 %**	0 %
UMP08	c.546G>A	WT	**98 %**	99 %	**NA**	-
		Exon 6 skipping (r.533_657del) ‡	**2.3 %**	1 %	**NA**	-
UMP10	p.Cys370Tyr	WT	**79 %**	100 %	**100 %**	100 %
		Exon 9 skipping (r.998_1109del)	**13 %**	0 %	**0 %**	0 %
		Exon 8+9 skipping (r.875_1109del)	**7 %**	0 %	**0 %**	0 %
UMP11	p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	WT	**86 %**	98 %	**87 %**	100 %
		p.Pro1079_Tyr1080insLeuGlnValAspProGluPro	**14 %**	2 %	**13 %**	0 %
UMP13	c.7082-2A>G	WT	**76 %**	100 %	**100 %**	100 %
		Activation of a cryptic site +7 nt in exon 42 (r.7082_7088del)	**24 %**	0 %	**0 %**	0 %
UMP14	c.7437G>A	WT	**58 %**	99 %	**NA**	-
		Exon 43 skipping (r.7288_7437)	**12 %**	0 %	**NA**	-
		Activation of a cryptic splice site + 146 nt in exon 43 (r.7434_7437del)	**30 %**	1 %	**NA**	-
UMP15	c.546G>A*	Exon 6 skipping(r.533_657del)	**3 %**	100 %	**11 %**	100 %
	c.7082-2A>G*	Activation of a cryptic site +7 nt in exon 42 (r.7082_7088del)	**6 %**	0 %	**0 %**	0 %
	c.8155+3G>C †	Exon 50 skipping (r.8116_8155del)	**95 %**	0 %	**95 %**	0 %

The values correspond to the percentage of the total reads aligned to the aberrant mRNA probes, setting as 100% the sum of reads aligned against the aberrant mRNA probes and reads aligned against the wild-type mRNA probes. NA indicates not available; and WT, wild-type. *Mutations in cis. †Mutations in trans. ‡Transcripts with low % of reads in leukocytes.

SUPPLEMENTAL FIGURES

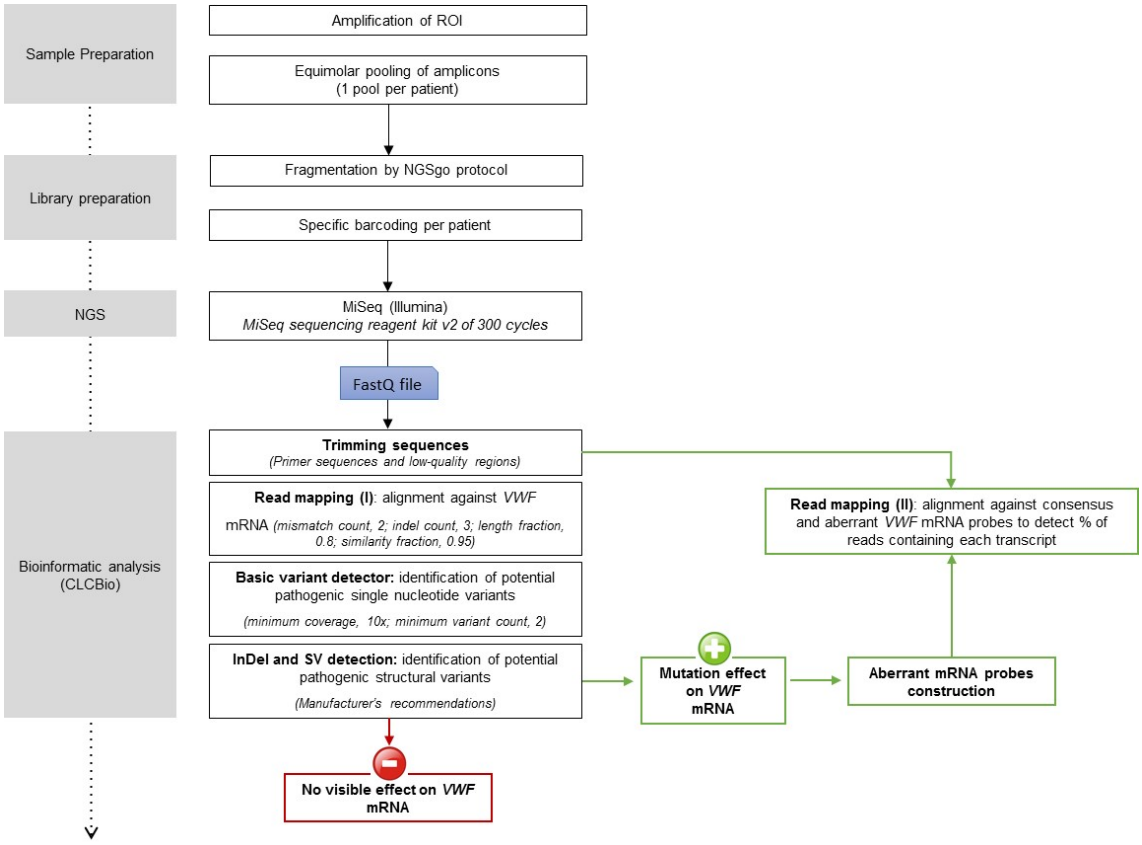


Figure S1. Flow chart used to analyze the effect of the mutation on mRNA using next-generation sequencing. ROI indicates regions of interest, NGS, next generation sequencing, and SV, structural variants.

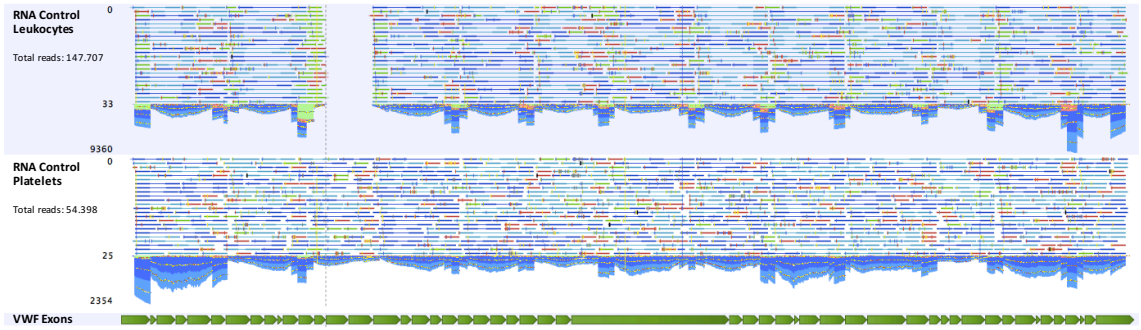


Figure S2. Schematic representation of alignment of NGS reads (from complete amplification of *VWF* mRNA from control platelets and leukocytes) to the *VWF* mRNA reference sequence. Because leukocytes show a predominant alternative splicing resulting in deletion of exon 14 (its start is shown with a discontinuous line) and exon 15, no reads were identified and mapped in this region. Except for exons 14 and 15 in leukocytes, all regions are presented with a minimum coverage of 10X in leukocytes and platelets.

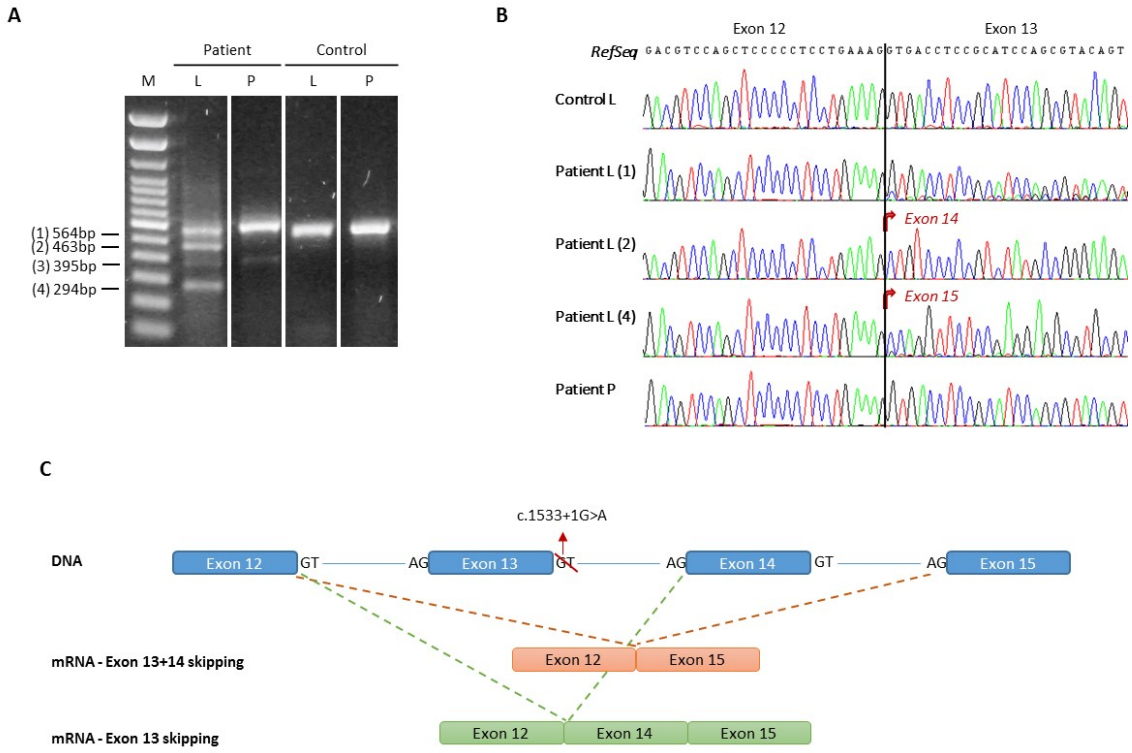


Figure S3. Analysis of the c.1533+1G>A mutation in patient UMP01. **A)** RT-PCR products amplified with primers located in exons 11 and 15 in RNA from leukocytes (L) and platelets (P) and separated on 1% agarose gel. Grouping of images from different parts of the same gel. **B)** Traditional Sanger sequencing of PCR product from patient L (1) showed the WT mRNA sequence, L (2) demonstrated exon 13 skipping and L (4) demonstrated exon 13 and 14 skipping. The L (3) band could not be purified in the agarose gel due to low concentration, thus it was not analyzed by Sanger. Platelets showed the WT mRNA sequence, indicating that the mutated allele had been degraded by NMD. **C)** Schematic representation of the mutation in genomic DNA and its effect on the *VWF* mRNA sequence. M indicates a 100-bp DNA ladder.

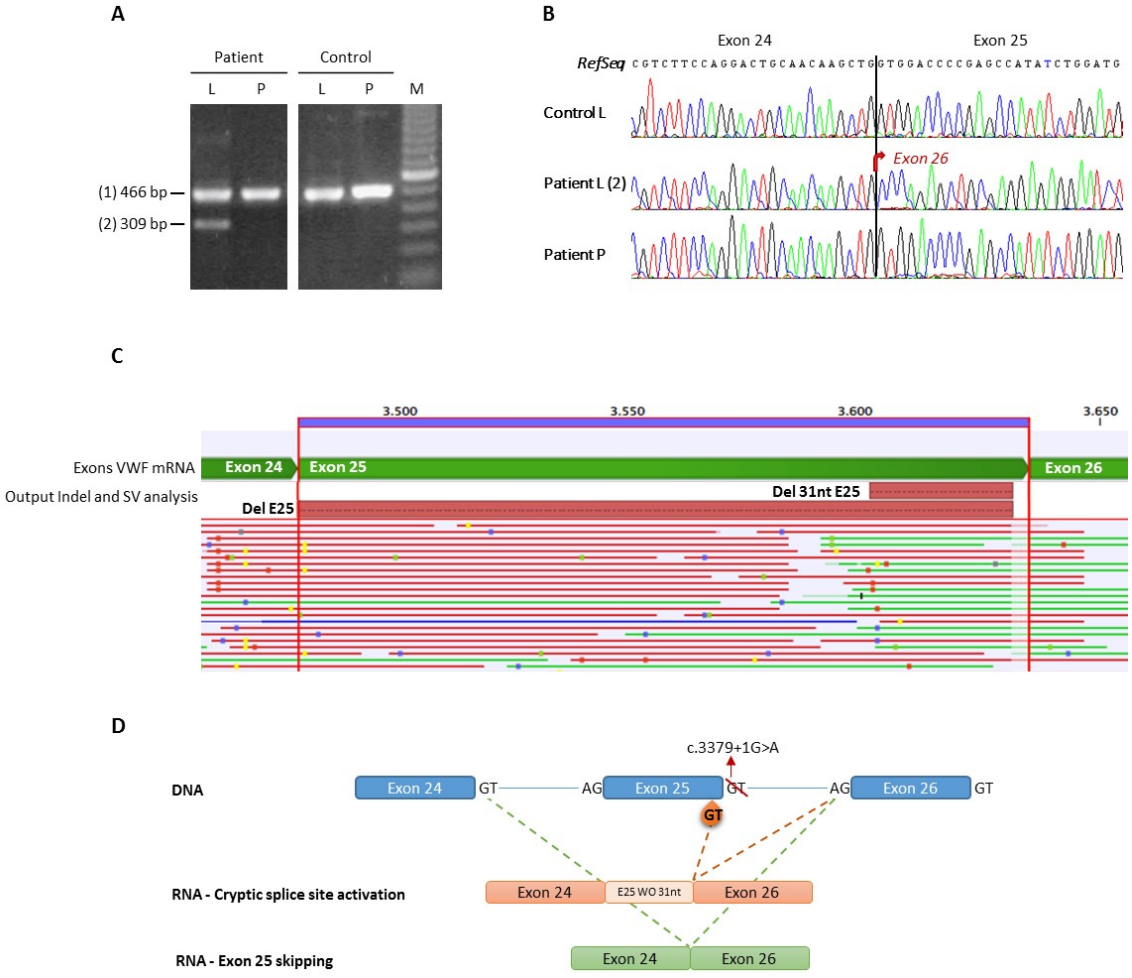


Figure S4. Analysis of the c.3379+1G>A mutation in patient UMP02. **A)** RT-PCR products amplified with primers located in exons 22 and 26 in RNA from leukocytes (L) and platelets (P) and separated on 1% agarose gel. Grouping of images from different parts of the same gel. **B)** Traditional Sanger sequencing of PCR product from patient L (2) showed skipping of exon 25. No changes were seen in the mRNA sequence in platelets, indicating that the mutated allele had been degraded by NMD. **C)** NGS of PCR products obtained from patient leukocytes identified two splicing variants: skipping of exon 25 and activation of a cryptic splice site—DSS 31 nt to WT-DSS—in exon 25. This last aberrant transcript was identified in a really low of transcripts. **D)** Schematic representation of the mutation in genomic DNA and its effect on the VWF mRNA sequence. M indicates 100-bp DNA ladder and; WO, without.

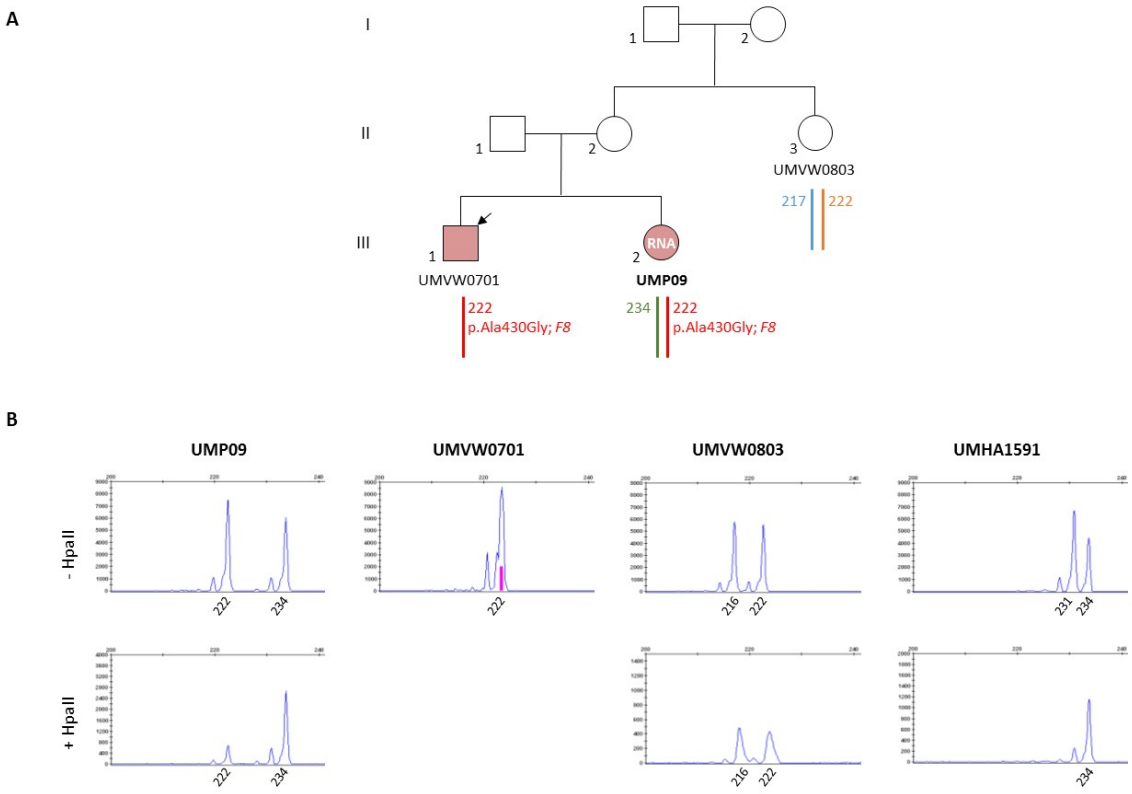


Figure S5. A) Pedigree of patient UMP09. The siblings, UMP09 and UMVW0701, showed the p.Ala430Gly mutation in the F8 gene in the X chromosome with the 222 peak. The arrow shows the index case. B) In patient UMP09, included in this study, her maternally-derived (UMVW0803) allele (222) was completely digested by the methylation-sensitive restriction enzyme Hpa II and, therefore, was not amplified by PCR. The remaining peak (234) is the paternally-derived allele representing the inactive, methylated X chromosome that resisted cleavage by Hpa II and was successfully amplified. Thus, this female patient shows complete skewing of X inactivation corresponding to the WT F8 gene. The UMHA1591 sample is a positive control for skewing of X inactivation.

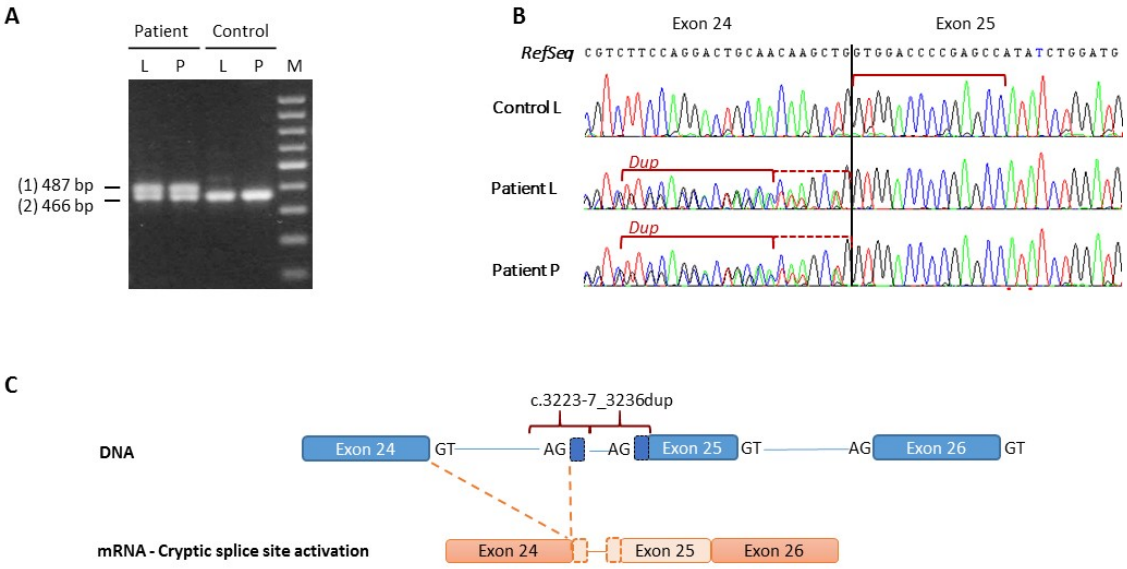


Figure S6. Analysis of the c.3223-7_3236dup mutation in patient UMP11. **A)** RT-PCR products amplified with primers located in exons 22 and 26 in RNA from leukocytes (L) and platelets (P) and separated on 1% agarose gel. **B)** Traditional Sanger sequencing of PCR product from patient L and P showed insertion of 21 nucleotides. Same results were observed by NGS (data not shown). **C)** Schematic representation of the mutation in genomic DNA and its effect on the VWF mRNA sequence. M, 100-bp DNA ladder.

REFERENCES

1. Corrales I, Ramirez L, Altisent C, Parra R, Vidal F. The study of the effect of splicing mutations in von Willebrand factor using RNA isolated from patients' platelets and leukocytes. J Thromb Haemost. 2011;9(4):679-688.

2. Favier R, Lavergne JM, Costa JM, et al. Unbalanced X-chromosome inactivation with a novel FVIII gene mutation resulting in severe hemophilia A in a female. Blood. 2000;96(13):4373-4375.

ARTICLE 5

Títol: "Caracterització de la variabilitat polimòrfica del gen del VWF en la cohort del projecte PCM-EVW-ES i influència en el fenotip."

Resum

El diagnòstic clínic de la VWD, especialment el tipus 1, pot ser complex a causa de diversos factors genètics i ambientals que influeixen en els nivells plasmàtics de VWF. S'estima que el 60% de la variació fenotípica en els nivells de VWF és atribuïble a factors hereditaris, amb un terç de la variància hereditària a causa de la influència del grup sanguini ABO. Tanmateix, alguns estudis recents proporcionen evidències de que els SNPs no sinònims poden contribuir de manera independent a la variabilitat fenotípica del VWF i del FVIII entre individus sans. Per tant, és probable que aquests factors contribueixin a la penetració incompleta i a la presentació clínica variable de la VWD. L'objectiu d'aquest treball és investigar la contribució potencial dels SNPs comuns del *VWF* al fenotip de la VWD considerant 219 pacients no relacionats del projecte Perfil Clínic i Molecular de la VWD a Espanya (PCM-EVW-ES). Es van combinar models de regressió lineal d'efectes mixts generalitzats per avaluar l'associació entre deu SNPs comuns i cinc mesures quantitatives relacionades amb la VWD: nivells d'antigen del VWF (VWF: Ag), activitat de cofactor de ristocetina (VWF: RCo), activitat d'unió al colágeno (VWF: CB), activitat de Factor FVIII (FVIII: C) i puntuació de sagnat. Aquesta estratègia ha permès descobrir l'associació entre les variants Thr789Ala, Val471Ile, Thr1381Ala i Gln852Arg VWF i les mesures de VWF: Ag, VWF: RCo i / o VWF: CB, que s'ha estimat que representen el 2,6% de la variància fenotípica. Els resultats obtinguts estan d'acord amb estudis anteriors i proporcionen proves addicionals per a la contribució potencial de les variacions comunes en la VWD. A més, aquests resultats poden servir de fonament per a futures anàlisi que condueixin a una millor comprensió de l'arquitectura genètica subjacent a la VWD.

Referència

Nina Borràs, Iris Garcia-Martínez, Javier Batlle, Almudena Pérez-Rodríguez, María Fernanda López-Fernández, Ángela Rodríguez-Trillo, Esther Lourés, Rafael Parra, Carme Altisent, Irene Corrales, and Francisco Vidal; on behalf of the members of the PCM-EVW-ES group.

Characterization of the polymorphic variability of the VWF gene in the PCM-EVW-ES registry cohort and influence on phenotype.

Journal of Thrombosis and Hemostasis (sotmès).

Characterization of the polymorphic variability of the VWF gene in the PCM-EVW-ES project cohort and influence on phenotype.

Nina Borràs,[1,2] Iris Garcia-Martínez,[1,2] Javier Batlle,[3] Almudena Pérez-Rodríguez,[3] Rafael Parra,[1,2] Carme Altisent,[2] María Fernanda López-Fernández,[3] Ángela Rodríguez-Trillo,[3] Esther Lourés,[3] and [1] Irene Corrales,[1,2] and Francisco Vidal;[1,2,5] in the name of all members of the PCM-EVW-ES group.

[1]Banc de Sang i Teixits, Barcelona, Spain; [2] Medicina Transfusional, Vall d'Hebron Research Institute, Universitat Autònoma de Barcelona (VHIR-UAB), Barcelona, Spain; [3]Complexo Hospitalario Universitario A Coruña, INIBIC, A Coruña, Spain;; and [4]CIBER de Enfermedades Cardiovasculares.

A complete list of the members of the PCM-EVW-ES appears as a supplemental data of this article.

Statement of equal authors' contribution: Nina Borràs and Iris Garcia-Martínez contributed equally to this work.

Running head: Contribution of *VWF* polymorphisms to VWD phenotype

Contact information for correspondence:
Francisco Vidal Pérez, PhD. Banc de Sang i Teixits, Barcelona, Spain. E-mail: **fvidal@bst.cat**

ABSTRACT

Clinical diagnosis of VWD, particularly type 1, can be complex because of several genetic and environmental factors that influence VWF plasma levels. It is estimated that 60% of the phenotypic variation in the VWF level is attributable to hereditary factors, with one-third of the hereditary variance due to the influence of the ABO blood group. However, some recent studies provide strong evidence that also nonsynonymous SNPs may independently contribute to VWF and FVIII phenotypic variability among healthy individuals. Thus, it is probable that these factors contribute to the incomplete penetrance and the variable clinical presentation of VWD. The aim of this work is to investigate the potential contribution of common SNPs within the *VWF* locus to the VWD phenotype considering 219 unrelated VWD patients from the Molecular and Clinical Profile of von Willebrand Disease in Spain project (PCM-EVW-ES). Generalized linear mixed-effects regression models were fitted to assess the association between ten common SNPs and five VWD-related quantitative measures: VWF antigen (VWF:Ag) levels, ristocetin cofactor activity (VWF:RCo), collagen binding activity (VWF:CB), Factor FVIII activity (FVIII:C) and Bleeding Score. This strategy allowed the uncovering of association between Thr789Ala, Val471Ile, Thr1381Ala and Gln852Arg *VWF* variants and VWF:Ag, VWF:RCo and/or VWF:CB measures, which were estimated to account up to 2.6% of the phenotypic variance. The obtained results are in accordance with previous studies and provide additional evidence for the potential contribution of common variations in VWD. In addition, these results may serve as a foundation for further analyses that lead to a better understanding of the genetic architecture underlying the VWD.

Key words: single nucleotide polymorphism, von Willebrand disease, von Willebrand factor

INTRODUCTION

The von Willebrand disease (VWD) is considered the most common congenital bleeding disorder found in the population and it is caused by a genetic defect in the von Willebrand factor gene (*VWF*) (1). The VWD is classified as a quantitative (types 1 and 3) or qualitative (type 2) defect of von Willebrand factor (VWF), a multimeric glycoprotein that plays an important role in primary hemostasis, promoting platelet adhesion at sites of vascular injury, as well as in secondary hemostasis serving as a carrier protein for Factor VIII (FVIII) (2). Clinical diagnosis of VWD, particularly type 1, can be complex because several genetic and environmental factors (e.g., blood group, age, sex, exercise, oral contraceptive use) can influence VWF plasma levels (3-5) and substantial variations are often evident during serial sampling in patients suspected of having the disease (6, 7). It is well established that about 60% of the phenotypic variation in the VWF levels is attributable to hereditary factors, with one-third of the hereditary variance due to the influence of the ABO blood group (8, 9). In type O individuals, the VWF level is 25%-35% lower than in non-O. Furthermore, genome-wide association studies were performed in recent years unraveling novel genetic loci influencing VWF:Ag levels: *STXBP5, SCARA5, STAB2, STX2, TC2N*, and *CLEC4M* (10). It is likely that variants on these receptors could also contribute to the incomplete penetrance and the variable clinical presentation of type 1 VWD.

Over the past years, enormous progress has been made in the understanding of the molecular basis of VWD and genetic studies have provided increasing evidence for the potential influence of common single nucleotide polymorphisms (SNPs) throughout the coding region of the *VWF* on the phenotypic variability of VWD (11). Although the role of common variation in this disorder has been incipient and poorly studied so far, the scenario is deeply changing due to advances brought about by the advent and improved implementation of Next Generation Sequencing (NGS) technologies, which are leading to a breakthrough in the genetics and molecular diagnosis of VWD. This technology eases the sequencing of the entire *VWF* and allows the identification of pathogenic and neutral sequence variants, in addition to those that have a subtle effect on VWF levels,

such as p.Thr789Ala (11). Moreover, the enforcement of genome and exome sequencing by NGS has enabled the identification of nearly 6,000 variants along the *VWF* coding region, yielding an average of one variant every 30 nucleotides. In line with this approach, a recent study analyzed the sequence of the *VWF* gene in 1,092 individuals who were part of the 1000 Genomes Project (12) (*http://www.internationalgenome.org/*) and identified over 2,728 polymorphisms as well as 91 insertions and deletions with a high level of ethnic bias, being the African population the most genetically variable, and showing the D' and D2 domains of *VWF* the highest level of diversity. Of the 94 non-synonymous variants identified, 31 were predicted to be deleterious, including 19 previously associated with VWD (13). Although most of these disease-associated variants display allele frequencies consistent with the reported incidence of the VWD in European cohorts and have been considered as causative, some of them show significantly higher frequencies in other ethnic groups. For instance, mutations p.Arg2185Gln, p.His817Gln and p.Met740Ile, usually associated with VWD type 1 and type 2N, were found in over 13% of the African sample (14). Coincidentally, a more recent study has demonstrated that some common SNPs reported in the healthy African-American (AA) population with known effects on VWF:Ag and FVIII:C parameters (i.e. p.Met740Ile and p.His817Gln), have a negligible frequency in the European population and, as a consequence, some of them have been traditionally considered responsible for the disease. In addition, this study provides strong evidence that at least six distinct non-synonymous *VWF* variants may independently contribute to VWF and FVIII phenotypic variability among healthy AA individuals, explaining the 3.3% and the 2.7% of the overall phenotypic variance, respectively (11). Finally, another study that considered healthy controls from the T. S. Zimmerman Program for the Molecular and Clinical Biology of VWD showed that for the AA controls, the presence of p.Ile1380Val, p.Asn1435Ser, and p.Asp1472His (all three SNPs in exon 28), was associated with a significantly lower VWF:RCo/VWF:Ag ratio; whereas the presence of p.Asp1472His alone was associated with a decreased ratio in both AA and Caucasian controls (15). In a posterior report, authors stated that this polymorphism was not associated with a significant increase in bleeding symptoms,

not even in type 1 VWD subjects (16).

Taking together all these data, new perspectives arise for the interpretation of common and rare SNPs, as well as the mutations widely accepted as responsible for the disease. This makes the mutational analysis of *VWF* (including SNPs) to emerge as an essential tool for the diagnosis and investigation of molecular etiology of the disease. However, until recently, molecular analyses of VWD were largely confined to merely identify the causative mutation without taking into account the polymorphisms' information. Furthermore, since the influence of common polymorphisms has been principally studied considering healthy cohorts and has been mainly focused in their effect on VWF levels, their influence remains to be more thoroughly explored in a phenotypically and genotypically well-characterized cohort of VWD patients. The possibility that their potential effects could explain some of the inter-individual and familial variability, and even help to delineate the blurred boundary that differentiates healthy individuals from patients with mild type 1 VWD, enhances the relevance of these lines of investigation.

In this regard, and based on the outcome of a Spanish survey highlighting the difficulties of diagnosing VWD (17), a multicenter prospective project (PCM-EVW-ES, Molecular and Clinical Profile of von Willebrand Disease in Spain) was designed to centrally characterize a large multicenter cohort of VWD patients, with the inclusion of the NGS analysis (18). The outcome of this study resulted in the molecular characterization of *VWF* in a clinical sample consisting of 556 individuals and the identification of a total of 704 variants (237 different) along *VWF* sequence (18, 19). It is important to highlight that the NGS approach implemented in the study provided reliable molecular data for cSNPs previously described in *VWF*, which have been further analyzed to determine their potential influence on the VWD phenotype, as has been done in healthy populations (11, 20).

From this perspective and to further disentangle the potential contribution of common SNPs within the *VWF* locus to the VWD phenotype, a comprehensive statistical analysis was performed considering the ten most frequent polymorphisms identified in the PCM-EVW-ES cohort and several disease-associated laboratory quantitative

parameters (Bleeding Score (BS), VWF:Ag, VWF:RCo, FVIII:C and VWF:CB). Generalized linear mixed-effects models were fitted to data considering the allelic dosages of the selected common variants in a PCM-EVW-ES subsample of 219 non-related subjects with VWD, and potential multiple-marker effects were further explored by conducting additive and epistatic analyses, as well as haplotype studies.

MATERIAL AND METHODS

Study Sample

Participants for the present study were recruited from the PCM-EVW-ES Project, which includes 480 VWD patients from 280 families, exhaustively characterized phenotypic and genotypically. The clinical diagnosis and inclusion criteria of the PCM-EVW-ES registry were previously described (19).

Exclusion criteria for the current study include known pregnancy and recent replacement therapy at the time of blood extraction, as well as missing information regarding variables with potential influence on the analysed phenotypes (ABO group, age and sex). In addition, since type 3 VWD is the most severe form of the disease, and is associated with a total or near-total absence of VWF, subjects with type 3 VWD diagnosis were not considered for the present analyses (21). Of the remaining sample, non-related subjects were selected, applying a random selection method for those families with more than one member.

The study was performed according to the guidelines of the Declaration of Helsinki, was approved by the local Research Ethics Committee, and all participants provided written informed consent.

Genotyping analysis by NGS

Molecular analysis of relevant regions of the *VWF*, including exons 1 to 52, intronic flanking regions, and approximately 1300bp of the promoter region, was performed by

means of a simple, fast and innovative protocol based on microfluidic technology and NGS as previously described (18, 19).

SNP and phenotypic variable selection

From the 188 different *VWF* variants identified in the molecular analysis described above, considering the 480 patients from the PCM-EVW-ES Spanish cohort, the 10 most frequent missense SNPs with a MAF>5% were selected to be further investigated. In regard to the VWD-related parameters to be analyzed in the present study, plasma measures of VWF:Ag, VWF:RCo, FVIII:C and VWF:CB, as well as the BS, were recorded from each patient to assess the contribution of common SNPs to the VWD phenotype. Complete laboratory measurements for each participant of the PCM-EVW-ES registry have been previously reported (19).

Analysis *in silico*

To evaluate the functional effects of each missense variants on VWF structure/function, *in silico* functional prediction was conducted by means of PolyPhen-2, the Sorting Intolerant From Tolerant (SIFT) and Mutation Taster (22-24).

Statistical Methods

Minimum allele frequencies (MAF) for each SNP, deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) patterns were evaluated considering the clinical sample of 219 subjects with VWD. Allele frequency comparisons were performed between the VWD sample, the five super-populations from the 1000 Genomes Project –African, Ad Mixed American, East Asian, European, South Asian– ((12); *http://www.internationalgenome.org/),* and the CIBERER cohort –excluding subpopulation III individuals, related to diseases of the blood and blood-forming organs (*http://www.ciberer.es/*; Collaborative Spanish Variant Server, *http://csvs.babelomics.org/*). Allele counts were calculated taking into account sample sizes and contrasts between proportions were assessed by Chi-Square tests.

Single-Marker Analyses, and Additive and Epistatic Effects. To assess the potential impact of common genetic variants on the five selected phenotypic variables, generalized linear mixed-effects regression models (GLMM) were fitted to data. Assuming an additive model of inheritance, SNP genotypes were coded as allelic dosages (0, 1, 2) of the alternative allele (Reference Genome, GRCh37/hg19; *http://genome.ucsc.edu/*). The medical recruitment unit of each subject (named "*centre*") was included in the models as a random-effect parameter. Sex, age and ABO group, dichotomically coded as type O or non-O, were considered as covariates, and thus as fixed-effect components, when a trend towards association with the outcomes was observed (P-value>0.20); and were kept in the fitted model if the corresponding regression coefficient was nominally significant (25). In this sense, while no association for sex was observed, age and ABO group showed suggestive or nominal association with all VWD-related measures and both were initially taken into account. However, when GLMMs were fitted, only the ABO group remained significant for the VWF:RCo and VWF:CB variables. Four quantitative traits (VWF:Ag, VWF:RCo, FVIII:C and VWF:CB) were continuous non-negative right-skewed data that displayed a good fit to the Gamma family distribution with a log link function (26, 27). Therefore, the exponential of β coefficients (exp(β)) derived from the GLMMs corresponds to the multiplicative effect on the expected quantitative trait for one-unit increase of the allelic dosage. For BS, a linear mixed-effect model was fitted, and β coefficients express the degree of change of the trait for every one-unit of change in the allelic dosage.

SNPs were declared nominally associated with the quantitative traits when the corresponding regression coefficient reached nominal significance. For nominally associated SNPs, a stepwise regression procedure was implemented to assess overall significance of fixed-effects, and GLMMs were compared by likelihood ratio tests to their correspondent intercept-only model. In addition, potential additive and epistatic effects between nominal associated SNPs were further explored applying the same procedure, inverting allelic dosages when tested SNPs displayed opposite effects in the single-marker analysis. Finally, the proportion of the total variance explained by fixed effects –marginal coefficient of determination R^2 was also calculated (28, 29).

All statistical tests described in this section were performed using the R environment and packages (*https://www.R-project.org/*; (30-32)).

Multiple-Marker Analyses. To avoid multiple testing and type I errors (α), we decided *a priori* to restrict the haplotype-based analyses to those SNPs showing nominal association with the outcomes in the single-marker study. The best two-marker haplotype from all possible combinations with frequencies over 0.05 was identified using the UNPHASED software, and when available, additional markers were added in a stepwise manner (v3.1.7; (33, 34)). Once identified the associated haplotypes, they were assigned to individuals with the PHASED 2.0 software (v2.1; (35, 36)), and GLMMs were fitted to evaluate the potential additive effects considering carriers of 0, 1 and 2 copies of each haplotype.

Nominal significance threshold for all tests was set at P-value<0.05. To address multiple-testing correction and give strong control of the family-wise error rate, the conservative Holm-Bonferroni method was firstly applied, considering the total number of hypothesis tested (37). It included statistical tests for each beta coefficient in full and final GLMMs, and those for the assessment of overall significance of GLMMs considering nominal significant SNPs (360 tests). However, since tests are not fully independent, and this method may entail certain loss of power and may be too stringent to identify subtle genetic effects, the False Discovery Rate (FDR) procedure with the Benjamini-Hochberg (BH) algorithm was additionally used to measure and control the proportion of false positives incurred in the study (38). Multiple-testing correction in haplotype analyses was conducted by permutation with the UNPHASED software, considering 10,000 permutations.

RESULTS

Patient selection and SNP cohort structure

Of the 480 subjects with VWD from the PCM-EVW-ES cohort, two were excluded due to pregnancy and recent replacement therapy, respectively. For the present study, 49 individuals with missing relevant data and/or 68 type 3 VWD diagnosis were neither considered. After applying exclusion criteria, a total of 219 non-related individuals with VWD were finally included in the analyses (Table 1, Figure 1).

Table 1. *Main demographic and clinical characteristics of the sample considered in the present study, consisting of 219 subjects with VWD recruited from the PCM-EVW-ES registry cohort.*

Characteristics		Type 1 VWD (N=93)	Type 2 VWD (N=118)	Unclassified (N=8)	All (N=219)
Female	N (%)	56 (60%)	67 (57%)	2 (25%)	90 (41%)
Male	N (%)	37 (40%)	51 (43%)	6 (75%)	129 (59%)
Blood group O	N (%)	60 (65%)	63 (53%)	8 (100%)	131 (60%)
Age	mean (SD)	40 (18.6)	40 (17.9)	35 (14.3)	40 (18.1)
Bleeding Score	mean (SD)	6 (5.4)	9 (5.7)	3 (3)	7 (5.8)
	median	4	9	3	7
VWF:Ag (UI/dL)	mean (SD)	26 (17.7)	38 (24.4)	54 (25.6)	34 (22.8)
	median	24	32	44	29
VWF:RCo (UI/dL)	mean (SD)	24 (15.8)	19 (21)	42 (27.8)	22 (19.7)
	median	21	9	44	16
VWF:CB (UI/dL)	mean (SD)	24 (14.8)	22 (18.9)	41 (22.5)	24 (23.7)
	median	23	16	39	19
FVIII:C (UI/dL)	mean (SD)	45 (24.6)	44 (24.6)	61 (20.5)	45 (24.6)
	median	46	36	55	43

SD indicates Standard Deviation and; N, Sample size

MAFs, LD patterns and P-values for deviation from HWE were calculated considering the sample of 219 subjects with VWD (Table 2). Although MAFs for all selected SNPs where above 0.05 in the PCM-EVW-ES cohort and displayed negligible differences compared to the final study sample of 219 individuals (data not shown), two variants, p.Gly2705Arg and p.Asn318Lys, showed frequencies slightly below 0.05 in the later. No significant departures from the HWE were detected for any of the SNPs (P-value<0.05) and LD patterns between them were weak (maximum r^2<0.3). The in-silico analysis

revealed that only two of them, p.Gly2705Arg and p.Phe2561Tyr, were predicted as deleterious by the PolyPhen2 and/or Sift algorithms (Table 2).

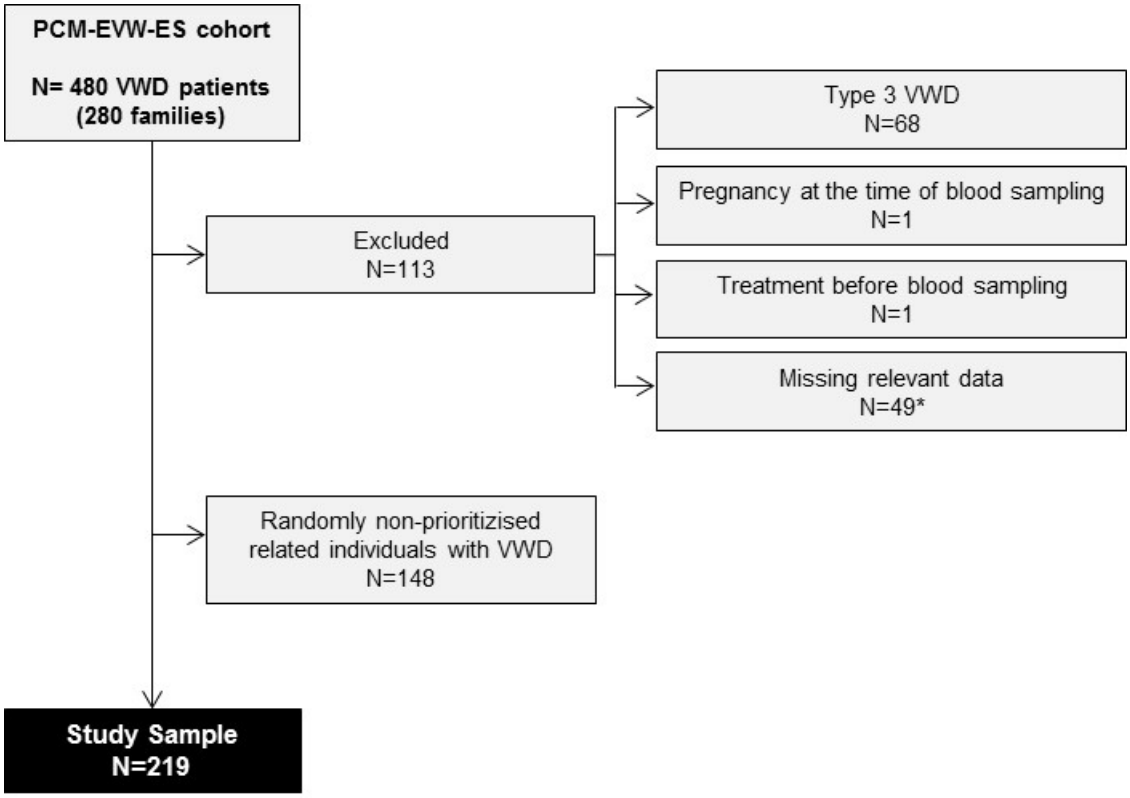


Figure 1. *Flowchart of the sample selection. VWD indicates von Willebrand disease; PCM-EVW-ES, Molecular and clinical profile of von Willebrand disease in Spain and; N, sample size.*Six patients*

When alternative allele frequencies were compared to different population cohorts, the African (AFR) and both Asian populations (EAS and SAS) showed significant divergence in all and in the 90% (9/10) of the SNPs, respectively; while for the admixed American group (AMR), significant differences were only detected for four SNPs (P-value<0.05). When compared to the European cohort, only the p.Thr1381Ala frequency slightly differed from the one reported in the VWD study sample (0.66 vs 0.6, P-value=0.041), and the p.Thr1381Ala and p.Asp1472His frequencies when compared to the CIBERER-CSVS cohort (0.105 vs 0.076, P-value=0.044; 0.66 vs 0.578, P-value=1.29e-03; Supplementary Table S1; Figure 2).

Table 2. Missense VWF common genetic variants identified in the whole PCM-EVW-ES cohort (N=480; MAF>0.05).

SNPdb ID	Ref Allele[a]	Alt Allele[a]	NT Change[a]	MAF	Genotype Counts	HWE P-value	AA Change[c]	Codon[c]	Exon	Domain	Sift	Polyphen	Mutation Taster
rs7962217	C	T[b]	c.8113G>A	0.048	1/19/199	0.395	p.Gly2705Arg	Gga/Aga	49	C6	deleterious (0)	possibly damaging (0.7)	polymorphism
rs35335161	A	T[b]	c.7682T>A	0.053	1/21/197	0.456	p.Phe2561Tyr	tTt/tAt	45	C4	tolerated (0.52)	possibly damaging (0.675)	polymorphism
rs1800385	C	A[b]	c.4693G>T	0.075	2/29/188	0.345	p.Val1565Leu	Gtg/Ttg	28	A2	tolerated (0.56)	benign (0.104)	polymorphism
rs1800383	C	G[b]	c.4414G>C	0.105	3/40/176	0.714	p.Asp1472His	Gac/Cac	28	A1	tolerated (0.15)	benign (0)	polymorphism
rs216311	T[b]	C	c.4141A>G	0.34	24/101/94	0.764	p.Thr1381Ala	Acc/Gcc	28	A1	tolerated (1)	benign (0.001)	polymorphism
rs216321	T[b]	C	c.2555A>G	0.068	2/26/191	0.261	p.Gln852Arg	cAg/cGg	20	D'	tolerated (1)	benign (0)	polymorphism
rs1063856	T	C[b]	c.2365A>G	0.347	30/92/97	0.297	p.Thr789Ala	Acc/Gcc	18	D'	tolerated (1)	benign (0.001)	polymorphism
rs1800378	T[b]	C	c.1451A>G	0.32	19/102/98	0.352	p.His484Arg	cAt/cGt	13	D2	tolerated (1)	benign (0.121)	polymorphism
rs1800377	C	T[b]	c.1411G>A	0.078	2/30/187	0.624	p.Val471Ile	Gtc/Atc	12	D2	tolerated (1)	benign (0.064)	polymorphism
rs1800387	A	T[b]	c.954T>A	0.048	2/17/200	0.076	p.Asn318Lys	aaT/aaA	8	D1	tolerated (0.3)	benign (0.023)	polymorphism

Data displayed in the table are based in the final clinical sample considered in the present study, consisting of 219 unrelated subjects with VWD. Results from the *in silico* functional analyses are shown. AA indicates Amino acid; CHR, Chromosome; HWE, Hardy-Weinberg Equilibrium; MAF, Minor Allele Frequency; N, sample size and; NT, nucleotide.

[a]Human Genome reference/alternative allele; GRCh37/hg19; http://genome.ucsc.edu/, [b]Minor Allele, [c]In the positive strand

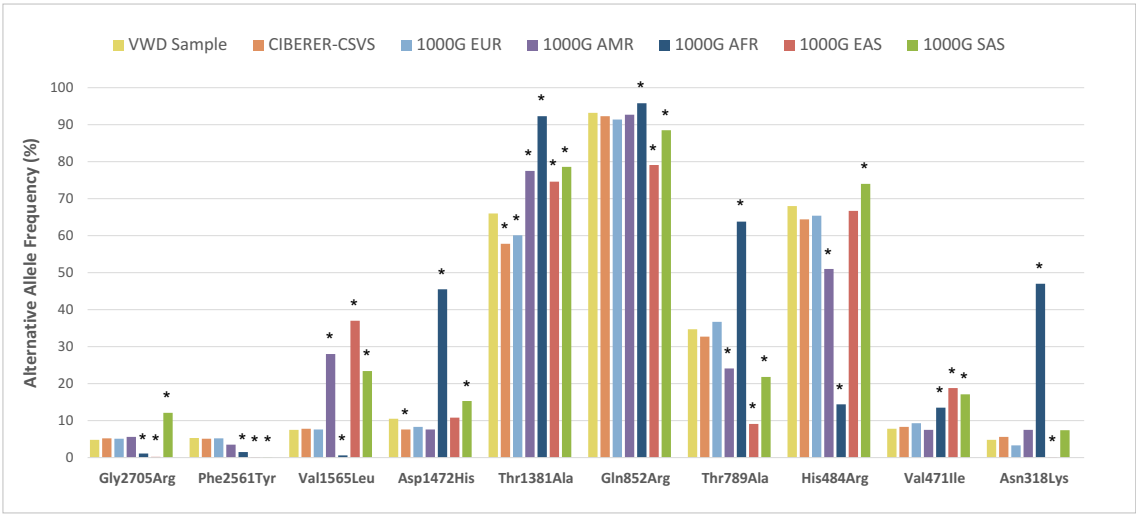


Figure 2. *Comparison of population allele frequencies of the ten missense VWF common genetic variants included in the present study. Y-axis represents alternative-allele frequencies in percentage and X-axis displays the seven compared populations grouped by each analysed common variant. Statistical significance of the difference between allelic proportions was assessed by Chi Square ($\chi 2$) test taking the VWD clinical sample as reference group. AFR indicates African; AMR, Admixed American; EAS, East Asian; SAS, South Asian; EUR, European; CIBERER-CSVS, Centro de Investigación Biomédica en Red - Collaborative Spanish Variant Server. *Nominal significance at P-value<0.05 from $\chi 2$ test for allelic proportions between the VWD study sample and population cohorts.*

Association study considering common missense SNPs in VWF

Single-Marker Analyses, and Additive and Epistatic Effects. Seven nominal associations were detected between p.Thr789Ala, p.Val471Ile, p.Thr1381Ala, and p.Gln852Arg genetic variants and VWF:Ag, VWF:RCo and/or VWF:CB measures (Table 3, Figure 3). No association was observed for neither FVIII:C nor BS parameters with any of the studied polymorphisms. Although none of the identified associations survived the Holm-Bonferroni correction for multiple comparisons (P-value$_{HB}$=0.05/360tests=1.39e-04), several remained significant after controlling the FDR at level of α = 5% (nominal P-value=0.022). The p.Thr789Ala alternative allele (C) dosage was found significantly associated with higher levels of VWF:Ag, showing an increase factor of exp(β_1) = 1.186, which corresponds to the ratio of predicted VWF:Ag levels for CC vs. TC, and TC vs. TT subjects. These results indicate that for each additional alternative allele (C), the VWF:Ag levels are estimated to be multiplied by 1.186, which represents an increase of approximately 19%. In other words, considering the mean VWF:Ag levels predicted for the p.Thr789Ala reference genotype (29.2 UI/dL for TT individuals), this measure is expected to be 1.186-fold higher for each additional copy

of the alternative allele (34.7 UI/dL (29.2*1.186) for TC and 41.2 UI/dL (34.7*1.186) for CC individuals). The p.Thr1381Ala polymorphism also displayed significant association with VWF:Ag, but in contrast to the later, the alternative allele (C) dosage was inversely correlated with VWF:Ag levels, with an estimated decrease factor of exp(β_1) = 0.855 per additional copy of the alternative allele (~14,5%). The proportion of the total variance of the VWF:Ag measure explained by the p.Thr789Ala and p.Thr1381Ala variants was estimated to be 2.6% and 2%, respectively. Additionally, the p.Gln852Arg alternative allele (C) dosage was also found significantly associated with decreasing levels of VWF:RCo, with a substantial reduction factor of exp(β_1) = 0.693 per allele (~30.7%). The proportion of the variability of the VWF:RCo measure explained by the p.Gln852Arg variant after controlling by the ABO group reached the 5.5%, of which, around 1.8% may be attributable to the effect of the SNP (data not shown). Direct nominal associations were detected between p.Val471Ile and VWF:Ag, and p.Thr789Ala and VWF:CB measures; and inverse nominal associations were identified between p.Thr1381Ala and VWF:RCo, and p.Gln852Arg and VWF:CB (Table 3, Figure 3).

When potential additive and epistatic effects were explored among nominally associated variants, a significant combined and independent effect for VWF:CB was detected between p.Thr789Ala and p.Gln852Arg variants, after controlling by a FDR of 5% (Table 4a). This analysis revealed that individuals with one p.Thr789Ala(C) allele and one p.Gln852Arg (T) allele are expected to show 1.692-fold higher VWF:CB levels compared to individuals with none of them, which translates into an increase of approximately 69%. The joined effect of these variants was estimated to account for 6.1% of the VWF:CB variance. In addition, a nominal significant epistatic effect regarding p.Thr789Ala and Val471Ile polymorphisms was identified for VWF:Ag levels, although the interaction term indicated a counteracting effect, which may be interpreted as a reduction of one SNP's estimated effect across levels of the other (Table 4b; Supplementary Figure S1). This model was estimated to explain 5.1% of the total VWF:Ag variance of which, around 3.8% may be attributed to the joint effect of these two variants.

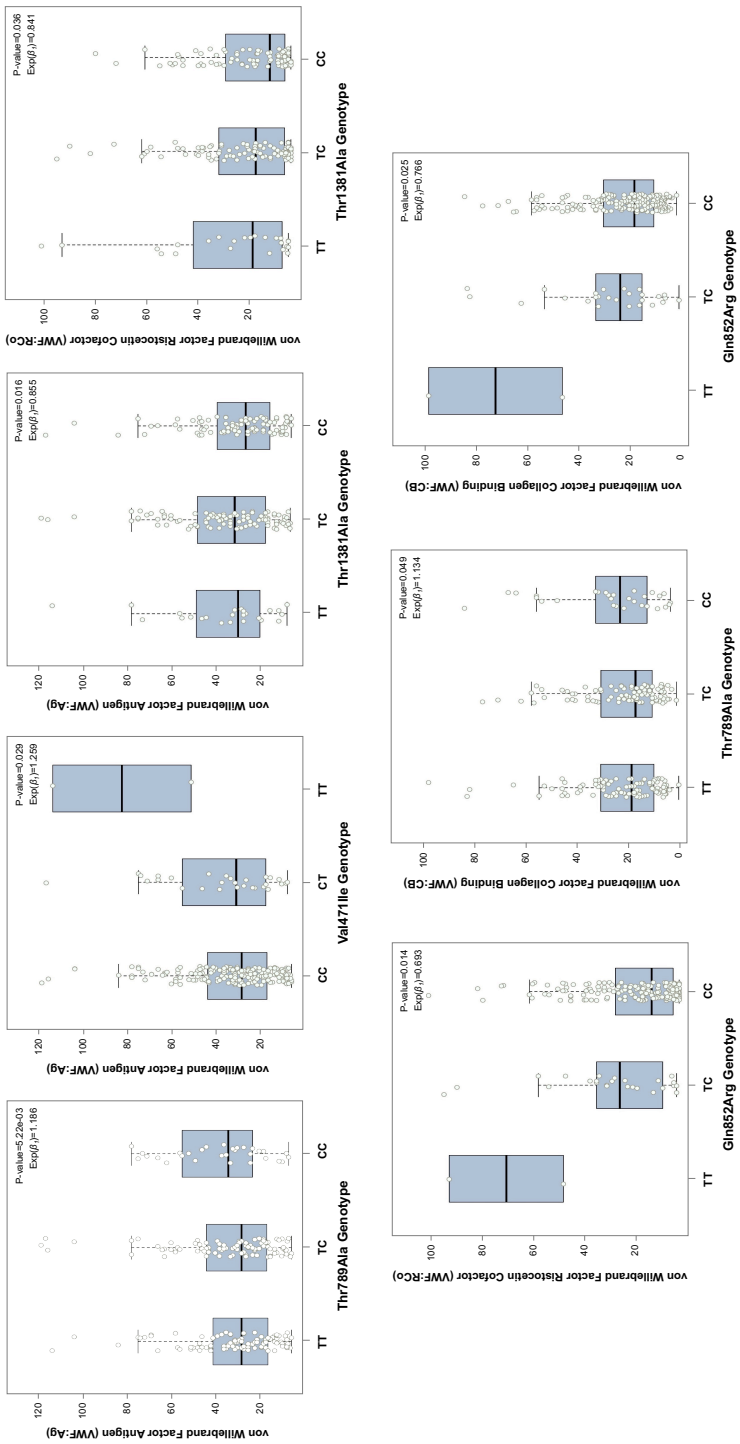


Figure 3. Boxplots of the association between VWD-related laboratory measures and genotypes for nominally associated SNPs in the VWF locus, considering 219 subjects with VWD. The Y-axes show the levels for VWF:Ag, VWF:RCo or VWF:CB parameters and the X-axes correspond to each genotype for p.Thr789Ala, p.Val471Ile, p.Thr1381Ala and p.Gln852Arg associated SNPs, displayed in increasing order of the alternate allele. Box midlines represent the median of the data, and the upper and lower limits of the box, the third and first quartile (75th and 25th percentiles) respectively. Whiskers extend up to 1.5 times the interquartile range from the top/bottom of the box to the furthest datum within that distance; any data beyond that distance, are represented individually as outlier points. Nominal P-values for regression beta coefficients for each marker derived from the GLMMs, as well as the exponential of betas (Exp(β_1)), are displayed for each model.

Table 3. Single-marker GLMM regression models considering ten SNPs and the phenotypic variables VWF:Ag, VWF:RCo and VWF:CB, in a clinical sample of 219 subjects with VWD.

VWF:Ag

			SNP		Covariates			
AA Change	**SNPdb ID**	**Alt Allele[a]**	**Exp(β_1)[b]**	**P-value (P-value$_{BH}$)**	**Exp(β_2)[b]**	**P-value (P-value$_{BH}$)**	**Marginal R^2**	**Model overall P-value (P-value$_{BH}$)**
p.Thr789Ala	rs1063856	C	1.186	**5.22e-03**** (0.013)	-	-	0.026	**5.54e-03**** (0,014)
p.Val471Ile	rs1800377	T	1.259	**0.029*** (0.056)	-	-	0.015	**0.025*** (0.051)
p.Thr1381Ala	rs216311	C	0.855	**0.016**** (0.034)	-	-	0.02	**0.016**** (0.034)

VWF:RCo

			SNP		Covariates: ABO			
AA Change	**SNPdb ID**	**Alt Allele[a]**	**Exp(β_1)[b]**	**P-value (P-value$_{BH}$)**	**Exp(β_2)[b]**	**P-value (P-value$_{BH}$)**	**Marginal R^2**	**Model overall P-value (P-value$_{BH}$)**
p.Thr1381Ala	rs216311	C	0.841	**0.036*** (0.068)	0.842	**9.1e-04**** (2.76e-03)	0.051	**1.17e-04***** (4e-04)
p.Gln852Arg	rs216321	C	0.693	**0.014**** (0.031)	1.475	**6.53e-04**** (2.03e-03)	0.055	**4.27e-05***** (1.47e-04)

VWF:CB

			SNP		Covariates: ABO			
AA Change	**SNPdb ID**	**Alt Allele[a]**	**Exp(β_1)[b]**	**P-value (P-value$_{BH}$)**	**Exp(β_2)[b]**	**P-value (P-value$_{BH}$)**	**Marginal R^2**	**Model overall P-value (P-value$_{BH}$)**
p.Thr789Ala	rs1063856	C	1.134	**0.049*** (0.092)	1.272	**0.011**** (0.026)	0.031	**7.54e-03**** (0.018)
p.Gln852Arg	rs216321	C	0.766	**0.025*** (0.05)	1.253	**0.017**** (0.036)	0.0363	**3.51e-03**** (8.8e-03)

Only SNPs showing nominal significant associations (P-value<0.05) are displayed. Neither FVIII:C nor BS results are shown since no SNP was detected displaying nominal significance. Covariates are only displayed when included in the final GLMM. Ag indicates Antigen; RCo, Ristocetin Cofactor; CB, Collagen Binding.*In bold, nominal significant results at nominal P-value<0.05. **Significant tests after BH multiple-testing correction considering a FDR$_{BH}$ of 5% and 360 tests; BH-adjusted P-value$_{BH}$ = 0.022.*** Significant tests after Holm-Bonferroni multiple-testing correction considering 360 tests; Holm-Bonferroni adjusted P-value = 1.39e-04. [a]Human Genome alternative allele; GRCh37/hg19; http://genome.ucsc.edu/. [b]Exp(β) indicates de multiplicative factor of each quantitative measure for one-unit increase of the allelic dosatge of tested SNPs.

Table 4. Analysis of (a) additive and (b) epistatic effects considering the nominally significant SNPs indentified in the single-marker studies. Only significant results are shown.

(**a**)

VWF:CB								
			SNP		**Covariates: ABO**			
AA Change	**SNPdb ID**	**Tested Allele**[a]	**Exp(β_1)**[b]	**P-value (P-value$_{BH}$)**	**Exp(β_2)**[b]	**P-value (P-value$_{BH}$)**	**Marginal R^2**	**Model overall P-value (P-value$_{BH}$)**
p.Thr789Ala	rs1063856	C	1.196	**6.22e-03**** (0.015)	1.263	**0.012**** (0.027)	0.061	**3.13e-04**** (1.06e-03)
p.Gln852Arg	rs216321	T	1.416	**3.79e-03**** (9.44e-03)				

(**b**)

VWF:Ag								
			SNP		**Covariates: ABO**			
AA Change	**SNPdb ID**	**Tested Allele**[a]	**Exp(β_1)**[b]	**P-value (P-value$_{BH}$)**	**Exp(β_2)**[b]	**P-value (P-value$_{BH}$)**	**Marginal R^2**	**Model overall P-value (P-value$_{BH}$)**
p.Thr789Ala	rs1063856	C	1.249	**1.47e-03**** (4.11e-03)	-	-	0.051	**1.56e-03**** (4.33e-03)
p.Val471Ile	rs1800377	T	1.484	**8.8e-03**** (0.021)	-	-		
p.Thr789Ala* p.Val471Ile			0.764	**0.042*** (0.079)	-	-		

Ag indicates Antigen; RCo, Ristocetin Cofactor and; CB, Collagen Binding.* In bold, nominal significance at P-value<0.05. ** Significant after BH multiple-testing correction considering a FDR$_{BH}$ of 5% and 360 tests; BH-adjusted P-value = 0.022. ***Significant after Holm-Bonferroni multiple-testing correction considering 360 tests; Holm-Bonferroni adjusted P-value = 1.39e-04. [a]For ease of interpretation, allelic dosages were inverted in this analysies when tested SNPs displayed opposite effects in the single-marker analysis. [b]Exp(β) indicates de multiplicative factor of each quantitative measure for one-unit increase of the allelic dosatge of tested SNPs.

***Multiple-Marker Analyses*.** The evaluation of the contribution of individual haplotypes to each VWD-related measure considering nominal associated SNPs revealed two combinations of markers significantly associated with the VWF:RCo (Global P-value=2.17e-03) or the VWF:CB (Global P-value=2.84e-03) parameters, after correcting by 10,000 permutations (Table 5a). The allelic combination p.Thr1381Ala(T)–p.Gln852Arg(T) was associated with higher levels of VWF:RCo, which is in accordance to the single-marker analysis results, where the dosage of the C allele of both variants was correlated to lower levels of this parameter (Table 3). The p.Thr789Ala(T)–p.Gln852Arg(T) combination was found significantly associated with higher levels of VWF:CB, although the interpretation of these effects is not as straightforward, since the p.Thr789Ala(C) allele dosage was the one found to be correlated with increasing levels of VWF:CB in the single-marker analysis (Table 3).

The additive effect of best haplotypes was further analysed by fitting GLMMs considering carriers of 0, 1 and 2 copies and both associations were confirmed (Table 5b). The p.Thr1381Ala(T)–p.Gln852Arg(T) additive dosage was significantly associated with increasing levels of VWF:RCo after controlling by a FDR of 5%, being estimated to be 1.421-fold higher for each additional copy of the haplotype, which corresponds to an increase of approximately the 42%. On the other hand, the p.Thr789Ala(T)–p.Gln852Arg(T) haplotype dosage was nominally associated with increasing levels of VWF:CB, being expected to be 1.305-fold higher per copy (~30%). The estimated phenotypic variance of the VWF:CB and VWF:RCo measures explained by the p.Thr1381Ala(T)–p.Gln852Arg(T) and p.Thr789Ala(T)–p.Gln852Arg(T) haplotypes was 5.3% and 3.6%, including the ABO group.

Table 5. Multiple-marker association results considering the SNPs surpassing the nominal significance threshold (P-value<0.05) in the single-marker analyses, considering the clinical sample of 219 subjects with VWD. Only nominally significant results are shown. (a) Haplotypes identified by the UNPHASED software; (b) GLMM regression results considering carriers of 0, 1 and 2 copies of the most significant haplotype for analysed VWD-related measure.

(**a**)

VWF: RCo						
Haplotype variants	**Haplotype alleles**	**Frequency (%)**	**Additive genetic value (CI 95%)[a]**	**Haplotype P-value**	**Overall association P-value**	**Empirical 5% quantile of the best P-value[b]**
p.Thr1381Ala – p.Gln852Arg	C - C	289 (66%)	0 (0 – 0)	0.061	**2.17e-03***	0.018
	T - C	119 (27%)	0.003 (-0.009 – 0.014)	0.83		
	T - T	30 (7%)	0.028 (0.013 – 0.043)	**9.62e-05*** (**1e-03**[+])		

VWF: RCo						
Haplotype variants	**Haplotype alleles**	**Frequency (%)**	**Additive genetic value (CI 95%)[a]**	**Haplotype P-value**	**Overall association P-value**	**Empirical 5% quantile of the best P-value[b]**
p.Thr789Ala – p.Gln852Arg	T - C	256 (58%)	0 (0 – 0)	**0.013***	**2.84e-03***	0.017
	C - C	152 (35%)	-0.003 (0.021 – 0.61)	0.435		
	T - T	30 (7%)	0.032 (0.014 – 0.049)	**7.28e-04*** (**3.9e-03**[+])		

(**b**)

VWF:RCo							
		SNP		**Covariates: ABO**			
Haplotype variants	**Haplotype alleles**	**Exp(β_1)[b]**	**P-value (P-value$_{BH}$)**	**Exp(β_2)[b]**	**P-value (P-value$_{BH}$)**	**Marginal R^2**	**Model overall P-value (P-value$_{BH}$)**
p.Thr1381Ala – p.Gln852Arg	T – T	1.421	**9.46e-03**** (0.022)	1.475	**3.37e-04**** (1.13e-03)	0.053	**4.08e-05***** (1.41e-04)

VWF:RCo							
		SNP		**Covariates: ABO**			
Haplotype variants	**Haplotype alleles**	**Exp(β_1)[b]**	**P-value (P-value$_{BH}$)**	**Exp(β_2)[b]**	**P-value (P-value$_{BH}$)**	**Marginal R^2**	**Model overall P-value (P-value$_{BH}$)**
p.Thr789Ala – p.Gln852Arg	T – T	1.305	**0.025*** (0.05)	1.253	**0.017**** (0.036)	0.036	**3.51e-03**** (8.8e-03)

Ag indicates Antigen; RCo, Ristocetin Cofactor and; CB, Collagen Binding. [a]Additive effects are shown relative to the reference haplotype. Additive value gives the change in expected trait value due to this haplotype, relative to the reference haplotype. [b]Fifth percentile point from the permutation distribution of the minimum P-value. It allows to assess whether other results are also significant after correcting for multiple testing. [c]Exp(β) indicates de multiplicative factor of the expected outcome (Y) for one-unit increase of the corresponding predictor (X). For small values of β (<0.2), 100*β approximates to the expected percentage change in Y for one-unit increase in X. *In bold, nominal significance at P-value<0.05. ** Significant after BH multiple-testing correction considering a FDR_{BH} of 5% and 360 tests; BH-adjusted P-value = 0.022. *** Significant after Holm-Bonferroni multiple-testing correction considering 360 tests; Holm-Bonferroni adjusted P-value = 1.39e-04. [+]Haplotype-based study best-P-value corrected by 10000 permutations using the UNPHASED software.

DISCUSSION

Under the hypothesis that, in addition to large-impact causative mutations, common variations may also contribute to the phenotype of VWD, the effects of ten frequent missense SNPs within *VWF* on several disease-related quantitative laboratory measures were explored in a clinical sample of 219 unrelated patients with VWD from the nation-wide PCM-EVW-ES cohort (19, 39). The results reported in the present study support this postulate and are in line with previous evidences that highlight the potential role of frequent polymorphisms in the disease. To our knowledge, although similar approaches have been applied considering healthy samples or general population, this is the largest study performed to date addressing the impact of *VWF* common genetic variation on several associated quantitative endophenotypes in a VWD sample. This approach has enabled the uncovering of association between p.Thr789Ala, p.Val471Ile, p.Thr1381Ala, and p.Gln852Arg variants and VWF:Ag, VWF:RCo and/or VWF:CB measures, which explained between 1.1 and 2.6% of the phenotypic variance of these measures. These variants are located in well-known functional *VWF* domains, named A1 (platelet-binding), D' and D2 (FVIII-binding and multimerization process) (Table 2), regions where the majority of previously described polymorphisms with reported influence on the VWD phenotype are located (14). For instance, a higher genetic variability has been described for D´ and D2 domains, which has been suggested to impact VWF multimeric patterns and VWF:Ag levels; and common polymorphisms in the A1 loop have been reported to affect ristocetin binding or ADAMTS13 cleavage efficiency (14, 15, 40, 41).

In the present study, the missense rs1063856 genetic variant, encoding for the p.Thr789Ala change, was found significantly associated with plasmatic VWF:Ag levels, and although some controversy exists in previous studies regarding the direction, most of them report higher levels of this parameter associated with the alternative allele (C) dosage, which is in accordance to the results provided here (10, 11, 42-47). This variant is located in the exon 18 of the *VWF*, within the FVIII-binding D' domain, which is critically involved in the VWF-FVIII interaction, as well as in normal multimerization and

optimal secretion of the VWF (11, 48, 49). *VWF* mutations in these domains lead to defects in the multimerization process and have been reported to result in intracellular retention and reduced secretion of VWF, and thus, in lower antigen levels in plasma (48, 50, 51). However, the p.Thr789Ala is a relatively conservative amino acid difference and is not predicted to be deleterious neither by Polyphen-2 nor Sift software (Table 2). In addition, no alteration in VWF expression levels nor in its interaction with FVIII has been previously observed *in vitro*, pointing that this polymorphism could be a marker for additional relevant variants exerting functional effects or could regulate VWF:Ag levels by a yet uncovered mechanism such as differential synthesis, processing and post-transcriptional modification, storage or plasmatic survival (47, 52). However, to disentangle whether the impact of p.Thr789Ala on VWF:Ag levels could be discriminated from the effects exerted by potential causative mutations, p.Thr789Ala effects were additionally tested controlling by the most frequent mutations detected in the VWD study sample (>1%). In this analysis, the association between p.Thr789Ala and VWF:Ag remained statistically significant, suggesting that the p.Thr789Ala effect on VWF levels was independent and could not be explained only by the presence of these mutations (data not shown). Interestingly, an epistatic effect was detected between this variant and the nominally associated rs1800377 polymorphism, which encodes for p.Val471Ile amino acid change, located in the D2 VWF domain. This region is involved in the multimerization of the VWF, and as described above, an impairment of this process may lead to decreased plasma levels of the antigen (48, 50, 51). The interaction between p.Thr789Ala and p.Val471Ile showed that their individual effects on VWF:Ag levels were conditioned on each other's genotype, resulting in a reduction of their impact on this parameter by each additional copy of the alternative allele (Supplementary Figure S1). Although these results should be taken with caution due to low genotypes frequency of p.Val471Ile in our VWD sample, this is the first time to our knowledge that tentative evidence for the contribution of SNP interactions in VWD is provided. Our findings suggest that further studies are warranted to define the role of epistatic effects in VWD and determine whether they may explain the incomplete penetrance and the variable clinical presentation of the disease.

On the other hand, in contrast to our findings, previous studies in large cohorts have described a significant association between the p.Thr789Ala variant and FVIII coagulation activity levels (11, 47). The lack of association in our study could be explained by demographic differences between samples, including ethnic, age, sex or ABO group distributions; sample size; or differential proportion of VWD patients or disease types. In relation to the later, VWF:Ag and FVIII:C measures were not fully correlated in our VWD sample (Spearman ρ = 0.63), with patients showing discordant levels of these parameters that may explain the inconsistencies.

In addition to these findings, a significant association was also uncovered between lower VWF:Ag levels and rs216311 alternative (C) allele dosage, which encodes for the p.Thr1381Ala change. Our results are in line with previous reports considering the Chinese population, although the directionality of the effects differ from the reported here, what may be attributable to the aforementioned reasons (53-55). In addition, the p.Thr1381Ala variant was also found nominally associated with decreasing levels of VWF:RCo activity, which evaluates the ability of VWF to agglutinate with normal platelets. The location of this variant in the A1 domain may explain this effect, since this region is involved in platelet GP1b receptor, heparin and possibly collagen binding (56).

Regarding VWF:RCo and VWF:CB, low levels of these parameters were found to correlate with the rs216321 alternative allele (C) dosage, coding for p.Gln852Arg change, located in the D' domain of the VWF (49). Previous studies had related this variant to alterations in VWF:Ag levels and/or FVIII:C coagulation activity, although none of them explored its impact on VWF:RCo or VWF:CB measures, which in contrast, reflect the functional ability of VWF to bind platelets or subendothelial collagen (43, 57). Additionally, in the haplotype analyses, both parameters were found associated with allelic combinations that included the rs216321 variant. However, since the TT genotype of this SNP showed a substantial low frequency and this may have impact the results, these findings should be taken as preliminary until further evidences supporting them are provided.

However, in spite of the evidences reported here supporting the involvement of common variation in VWD-related phenotypic measures, none of them was found associated to the BS (data not shown). Previous studies with similar findings have suggested that the variants would not be directly affecting bleeding symptoms (47). In contrast, we argue that, since different causes may lead to a specific score, an inherent genetic heterogeneity underlies the BS, which makes the unravelling of significant associations more difficult than considering VWD-related measurable traits that may act as endophenotypes for the disorder, with a more defined genetic architecture.

Several methodological considerations should be noted for the current study. First, heterogeneity in sample processing or laboratory assays, as well as unknown or not reported influential conditions of the participants –such as pregnancy or recent replacement therapy–, may have introduced certain degree of bias and lowered our statistical power to uncover genetic associations of subtle effects. Second, and in addition to the previous point, the modest size of the VWD sample may have as well prevented us from detecting significant associations of smaller magnitude, including previously reported variants, such as p.Val1565Leu or the p.Asp1472His (58). In addition, due to low genotype frequencies, some results should be interpreted cautiously and be considered preliminary until further replication in additional cohorts. In this regard, family-based study designs considering larger and exhaustively phenotyped samples, such as the whole PCM-EVW-ES registry cohort, will improve the statistical power of these approaches and lead to the identification of new implicated *loci* and higher replication rates of previous results. Third, extending the study to additional VWD-related measures, such as multimer analysis, and other critical genetic regions of the *VWF*, such as the promoter, will provide further relevant data to explore and, eventually, disentangle the complex genetic background underlying the VWD. Fourth, although allele frequencies for the studied variants have been compared between several populations cohorts, comprehensive case-control studies are needed to determine the potential involvement of these variants in the risk of developing VWD and identify the missing heritability.

To conclude, the exhaustive analysis presented here considering the most frequent polymorphic variants from the PCM-EVW-ES cohort and VWD-related phenotypic measures, has enabled us to provide new, as well as additional evidence for the contribution of common variation to the disease. Our findings highlight the importance of exploring this field and projects the basis for further analyses that lead to a better understanding of the complex genetic architecture underlying the VWD.

REFERENCES

1. Rodeghiero F, Castaman G, Dini E. Epidemiological investigation of the prevalence of von Willebrand's disease. Blood. 1987;69(2):454-9.

2. Sadler JE. Biochemistry and genetics of von Willebrand factor. Annu Rev Biochem. 1998;67:395-424.

3. Zhang Z, Lindstedt M, Blomb‰ck M, Anvret M. Effects of the mutant von Willebrand factor gene in von Willebrand disease. Hum Genet. 1995;96(4):388-94.

4. Castaman G, Eikenboom JC, Bertina RM, Rodeghiero F. Inconsistency of association between type 1 von Willebrand disease phenotype and genotype in families identified in an epidemiological investigation. Thromb Haemost. 1999;82(3):1065-70.

5. Wahlberg TB. A method for the evaluation of clinical information, exemplified for bleeding symptoms in non-severe von Willebrand's disease type I. Methods Inf Med. 1984;23(3):143-6.

6. Abildgaard CF, Suzuki Z, Harrison J, Jefcoat K, Zimmerman TS. Serial studies in von Willebrand's disease: variability versus "variants". Blood. 1980;56(4):712-6.

7. Zhang ZP, Deng LP, Blomback M, Anvret M. Dinucleotide repeat polymorphism in the promoter region of the human von Willebrand factor gene (vWF gene). Hum Mol Genet. 1992;1(9):780.

8. Gill JC, Endres-Brooks J, Bauer PJ, Marks WJJ, Montgomery RR. The effect of ABO blood group on the diagnosis of von Willebrand disease. Blood. 1987;69(6):1691-5.

9. Orstavik KH, Magnus P, Reisner H, Berg K, Graham JB, Nance W. Factor VIII and factor IX in a twin population. Evidence for a major effect of ABO locus on factor VIII level. Am J Hum Genet. 1985;37(1):89-101.

10. Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. Circulation. 2010;121(12):1382-92.

11. Johnsen JM, Auer PL, Morrison AC, Jiao S, Wei P, Haessler J, et al. Common and rare von Willebrand factor (VWF) coding variants, VWF levels, and factor VIII levels in African Americans: the NHLBI Exome Sequencing Project. Blood. 2013;122(4):590-7.

12. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.

13. The Genomes Project C. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.

14. Wang QY, Song J, Gibbs RA, Boerwinkle E, Dong JF, Yu FL. Characterizing polymorphisms and allelic diversity of von Willebrand factor gene in the 1000 Genomes. J Thromb Haemost. 2013;11(2):261-9.

15. Flood VH, Gill JC, Morateck PA, Christopherson PA, Friedman KD, Haberichter SL, et al. Common VWF exon 28 polymorphisms in African Americans affecting the VWF activity assay by ristocetin cofactor. Blood. 2010;116(2):280-6.

16. Flood VH, Friedman KD, Gill JC, Haberichter SL, Christopherson PA, Branchford BR, et al. No increase in bleeding identified in type 1 VWD subjects with D1472H sequence variation. Blood. 2013;121(18):3742-4.

17. Batlle J, Perez-Rodriguez A, Pinto JC, Fraga EL, Rodriguez-Trillo Tch A, Fernanda Lopez-Fernandez M. Diagnosis and management of von Willebrand disease in Spain. Semin Thromb Hemost. 2011;37(5):503-10.

18. Batlle J, Perez-Rodriguez A, Corrales I, Lopez-Fernandez MF, Rodriguez-Trillo A, Loures E, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): Proposal for a new diagnostic paradigm. Thromb Haemost. 2016;115(1):40-50.

19. Borras N, Batlle J, Perez-Rodriguez A, Lopez-Fernandez MF, Rodriguez-Trillo A, Loures E, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): comprehensive genetic analysis by next-generation sequencing of 480 patients. Haematologica. 2017;102(12):2005-14.

20. Bellissimo DB, Christopherson PA, Flood VH, Gill JC, Friedman KD, Haberichter SL, et al. VWF mutations and new sequence variations identified in healthy controls are more frequent in the African-American population. Blood. 2012;119(9):2135-40.

21. Ng C, Motto DG, Di Paola J. Diagnostic approach to von Willebrand disease. Blood. 2015;125(13):2029-37.

22. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Current protocols in human genetics. 2013;Chapter 7:Unit7 20.

23. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073-81.

24. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11(4):361-2.

25. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. Am J Epidemiol. 1993;138(11):923-36.

26. Burnham K, David R. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, editor2002.

27. Burnham K, Anderson D. Model Selection and Inference: A Practical Information-Theoretic Approach. 2nd Edition ed. Springer-Verlag, editor2002.

28. Johnson PC. Extension of Nakagawa & Schielzeth's R(2)GLMM to random slopes models. Methods in ecology and evolution. 2014;5(9):944-6.

29. Nakagawa S, Johnson PCD, Schielzeth H. The coefficient of determination R(2) and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of the Royal Society, Interface. 2017;14(134).

30. Lefcheck J. PiecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. Methods in Ecology and Evolution 2016;7(5):573-9.

31. Bates D, Maechler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 2015;67(1):1-48.

32. Delignette-Muller M, Dutang C. fitdistrplus: An R Package for Fitting Distributions. . Journal of Statistical Software 2015;64(4):1-34.

33. Dion I, Lahaye M, Salmon R, Baquey C, Monties JR, Havlik P. Blood haemolysis by ceramics. Biomaterials. 1993;14(2):107-10.

34. Fallin D, Schork NJ. Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. Am J Hum Genet. 2000;67(4):947-59.

35. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 2001;68(4):978-89.

36. Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet. 2005;76(3):449-62.

37. Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 1979;6:65–70.

38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B. 1995;57:289–300.

39. Batlle J, Perez-Rodriguez A, Corrales I, Lopez-Fernandez MF, Rodriguez-Trillo A, Loures E, et al. Molecular and clinical profile of von Willebrand disease in Spain (PCM-EVW-ES): Proposal for a new diagnostic paradigm. Thromb Haemost. 2016;114(6).

40. Campos M, Sun W, Yu F, Barbalic M, Tang W, Chambless LE, et al. Genetic determinants of plasma von Willebrand factor antigen levels: a target gene SNP and haplotype analysis of ARIC cohort. Blood. 2011;117(19):5224-30.

41. Pruss CM, Notley CR, Hegadorn CA, O'Brien LA, Lillicrap D. ADAMTS13 cleavage efficiency is altered by mutagenic and, to a lesser extent, polymorphic sequence changes in the A1 and A2 domains of von Willebrand factor. Br J Haematol. 2008;143(4):552-8.

42. Ahmad F, Kannan M, Biswas A, Saxena R. Impact of 789Ala/Ala genotype on quantitative type of von Willebrand disease. Ann Hematol. 2009;88(5):479-83.

43. van Loon JE, Kavousi M, Leebeek FW, Felix JF, Hofman A, Witteman JC, et al. von Willebrand factor plasma levels, genetic variations and coronary heart disease in an older population. J Thromb Haemost. 2012;10(7):1262-9.

44. Lacquemant C, Gaucher C, Delorme C, Chatellier G, Gallois Y, Rodier M, et al. Association between high von willebrand factor levels and the Thr789Ala vWF gene polymorphism but not with nephropathy in type I diabetes. The GENEDIAB Study Group and the DESIR Study Group. Kidney Int. 2000;57(4):1437-43.

45. Klemm T, Mehnert AK, Siegemund A, Wiesner TD, Gelbrich G, Bluher M, et al. Impact of the Thr789Ala variant of the von Willebrand factor levels, on ristocetin co-factor and collagen binding capacity and its association with coronary heart disease in patients with diabetes mellitus type 2. Exp Clin Endocrinol Diabetes. 2005;113(10):568-72.

46. Zabaneh D, Gaunt TR, Kumari M, Drenos F, Shah S, Berry D, et al. Genetic variants associated with Von Willebrand factor levels in healthy men and women

identified using the HumanCVD BeadChip. Annals of human genetics. 2011;75(4):456-67.

47. Flood VH, Johnsen JM, Kochelek C, Slobodianuk TL, Christopherson PA, Haberichter SL, et al. Common VWF sequence variants associated with higher VWF and FVIII are less frequent in subjects diagnosed with type 1 VWD. Res Pract Thromb Haemost. 2018;2(2):390-8.

48. Jorieux S, Fressinaud E, Goudemand J, Gaucher C, Meyer D, Mazurier C. Conformational changes in the D' domain of von Willebrand factor induced by CYS 25 and CYS 95 mutations lead to factor VIII binding defect and multimeric impairment. Blood. 2000;95(10):3139-45.

49. Rauch A, Wohner N, Christophe OD, Denis CV, Susen S, Lenting PJ. On the versatility of von Willebrand factor. Mediterranean journal of hematology and infectious diseases. 2013;5(1):e2013046.

50. Haberichter SL, Allmann AM, Jozwiak MA, Montgomery RR, Gill JC. Genetic alteration of the D2 domain abolishes von Willebrand factor multimerization and trafficking into storage. J Thromb Haemost. 2009;7(4):641-50.

51. Tjernberg P, Vos HL, Castaman G, Bertina RM, Eikenboom JC. Dimerization and multimerization defects of von Willebrand factor due to mutated cysteine residues. J Thromb Haemost. 2004;2(2):257-65.

52. Kroner PA, Foster PA, Fahs SA, Montgomery RR. The defective interaction between von Willebrand factor and factor VIII in a patient with type 1 von Willebrand disease is caused by substitution of Arg19 and His54 in mature von Willebrand factor. Blood. 1996;87(3):1013-21.

53. Fu Y, Zhu P, Bu DF, Zhang Y, Wang MY, Gong YW, et al. [A1381T and -1793G/C polymorphisms of vWF gene impact the plasma vWF levels in Yugur, Tibetan and Han nationalities of China]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2012;20(5):1200-4.

54. Yuan ZH, Hou YJ, Li Y, Zhang H, Li ZY, Li X, et al. [Analysis of vWF gene A1381T polymorphism in patients with coronary heart disease]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2011;19(3):775-80.

55. Yuan ZH, Zhao J, Zhang Y, Zhu P. [Impact of vWF gene A1381T polymorphism and ABO blood group on von Willebrand factor level in plasma]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2010;18(4):967-71.

56. Bharati KP, Prashanth UR. Von Willebrand disease: an overview. Indian journal of pharmaceutical sciences. 2011;73(1):7-16.

57. Tang W, Cushman M, Green D, Rich SS, Lange LA, Yang Q, et al. Gene-centric approach identifies new and known loci for FVIII activity and VWF antigen levels in European Americans and African Americans. Am J Hematol. 2015;90(6):534-40.

58. Davies JA, Bowen DJ. The association between the L1565 variant of von Willebrand factor and susceptibility to proteolysis by ADAMTS13. Haematologica. 2007;92(2):240-3.

Material Suplementari Article 5

Table S1. Comparisons of population allele frequencies of the ten missense *VWF* common genetic variants included in the present study. Significance for differences in allelic proportions between the VWD study sample and population cohorts is addressed by Chi Square (χ^2) tests.

		Gly2705Arg		Phe2561Tyr		Val1565Leu		Asp1472His		Thr1381Ala		Gln852Arg		Thr789Ala		His484Arg		Val471Ile		Asn318Lys	
Population	Sample size	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value	Alt allele counts (freq)[a]	χ^2 P-value
1000G AFR	661	14 (0.011)	**3.23e-06***	20 (0.015)	**2.52e-05***	8 (0.006)	**3.66e-16***	602 (0.455)	**<2e-16***	1220 (0.923)	**<2e-16***	1267 (0.958)	**0.032***	843 (0.638)	**<2e-16***	190 (0.144)	**<2e-16***	178 (0.135)	**1.98e-03***	622 (0.47)	**<2e-16***
1000G AMR	347	39 (0.056)	0.64	24 (0.035)	0.187	194 (0.28)	**<2e-16***	53 (0.076)	0.12	538 (0.775)	**2.75e-05***	643 (0.927)	0.842	167 (0.241)	**1.4e-04***	354 (0.51)	**2.34e-08***	52 (0.075)	0.959	51 (0.073)	0.112
1000G EAS	504	2 (0.002)	**6e-10***	1 (0.001)	**8.95e-12***	373 (0.37)	**<2e-16***	109 (0.108)	0.934	752 (0.746)	**9.98e-04***	797 (0.791)	**6.74e-11***	92 (0.091)	**<2e-16***	672 (0.667)	0.654	190 (0.188)	**1.33e-07***	1 (0.001)	**9.87e-11***
1000G SAS	489	118 (0.121)	**3.27e-05***	1 (0.001)	**1.88e-11***	229 (0.234)	**1.94e-12***	150 (0.153)	**0.019***	769 (0.786)	**5.87e-07***	866 (0.885)	**0.01***	213 (0.218)	**3.9e-07***	724 (0.74)	**0.0237***	167 (0.171)	**5.14e-06***	72 (0.074)	0.092
1000G EUR	503	51 (0.051)	0.929	52 (0.052)	1	76 (0.076)	1	83 (0.083)	0.201	605 (0.601)	**0.041***	919 (0.914)	0.295	369 (0.367)	0.501	658 (0.654)	0.3626	94 (0.093)	0.384	33 (0.033)	0.214
CIBERER-CSVS Cohort[b]	1612	168 (0.052)	0.799	164 (0.051)	0.975	251 (0.078)	0.929	245 (0.076)	**0.044***	1863 (0.578)	**1.29e-03***	2976 (0.923)	0.597	1054 (0.327)	0.432	2076 (0.644)	0.1483	268 (0.083)	0.7641	181 (0.056)	0.553
VWD Study Sample	219	21 (0.048)	-	23 (0.053)	-	33 (0.075)	-	46 (0.105)	-	289 (0.66)	-	408 (0.932)	-	152 (0.347)	-	298 (0.68)	-	34 (0.078)		21 (0.048)	-

Abbreviations: AFR = African; AMR = Admixed American; EAS = East Asian; SAS = South Asian; EUR = European; CIBERER-CSVS = Centro de Investigación Biomédica en Red - Collaborative Spanish Variant Server. [a]Total *alternative allele counts* are calculated taking into account alternative allele frequencies (Human Genome alternative allele; GRCh37/hg19; http://genome.ucsc.edu/) and sample sizes (N), applying the following formula: allele frequency*N*2. [b]Excluding subpopulation III individuals -*Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism*-. Partially overlaps with 1000G European (EUR) super-population panel (107 subjects from the Spanish subpopulation (IBS)).* In bold, nominally significant comparisons at P-value<0.05 from Chi Square (χ^2) test for allelic proportions between the VWD Study Sample and population cohorts.

SUPPLEMENTAL FIGURE

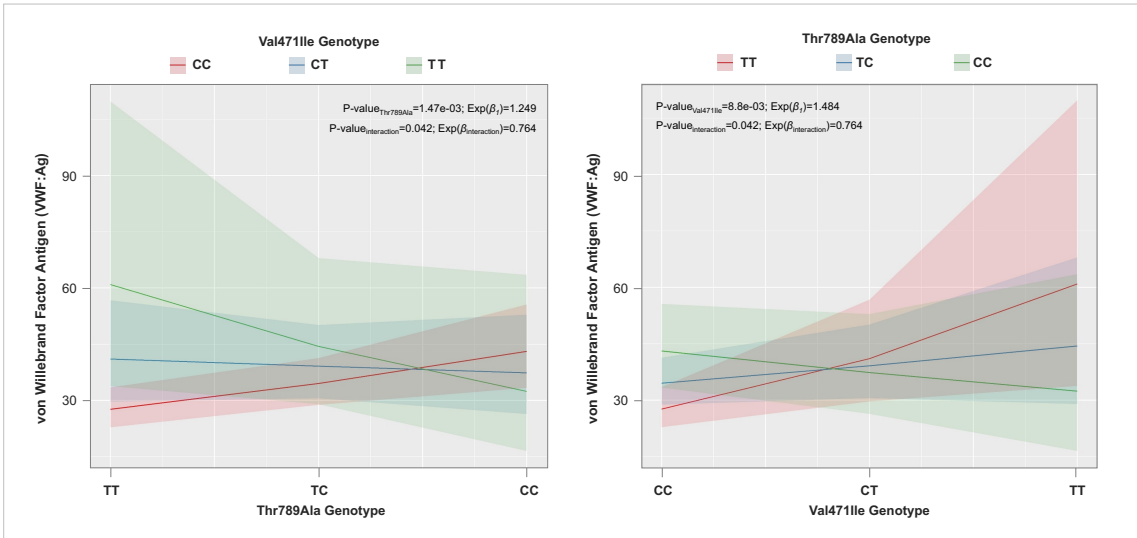


Figure S1. Interaction plot showing the epistatic effects between Thr789Ala and Val471Ile variants on VWF:Ag parameter. The Y-axis shows VWF:Ag levels and the X-axes correspond to the Thr789Ala and Val471Ile genotypes, displayed in increasing order of the alternate alleles. Each line represents the change in the estimated effect of one variant on VWF:Ag across the genotypes of the other variant. Nominal P-values for regression beta coefficients for each marker and interaction term derived from the GLMMs, as well as the exponential of betas (Exp(β)), are displayed in the plots.

LIST OF THE MEMBERS OF THE PCM-EVW-ES

Ana Rosa Cid, MD (S. Hematología, H.Universitario y Politécnico La Fe, Valencia, Spain); Santiago Bonanad, MD (S. Hematología, H. Universitario y Politécnico La Fe, Valencia, Spain); María Eva Mingot-Castellano, MD (S. Hematología, H. Regional Universitario de Málaga, Málaga, Spain); Nira Navarro, MD (S. Hematología, H. Universitario Dr. Negrín, Las Palmas de Gran Canaria, Spain); Rocío Pérez-Montes, MD (S. Hematología, H.Universitario Marqués de Valdecilla, Santander, Spain); Shally Marcellini, MD (S. Hematología, Salud Castilla y León, Segovia, Spain); Ana Moretó, MD (S. Hematología, H. Universitario Cruces, Barakaldo, Spain); Sonia Herrero, MD (S. Hematología, H. Universitario de Guadalajara, Guadalajara, Spain); Inmaculada Soto, MD (S. Hematología, H. Universitario Central de Asturias, Oviedo, Spain); Angel Bernardo-Gutiérrez, MD (S. Hematología, H. Universitario Central de Asturias, Oviedo, Spain); Nuria Fernández Mosteirín, MD (S. Hematología, H. Universitario Miguel Servet, Zaragoza, Spain); Víctor Jiménez-Yuste, MD (S. Hematología, H. Universitario La Paz, Madrid, Spain); Nieves Alonso, MD (S. Hematología, H. Infanta Cristina, Badajoz, Spain); Aurora de Andrés Jacob, MD (S. Hematología, Complexo Hospitalario Universitario Santiago de Compostela, Spain); Emilia Fontanes, MD (S. Hematología, H. Universitario Lucus Augusti, Lugo, Spain); Rosa Campos, MD (S. Hematología, H. Jerez de la Frontera, Cádiz, Spain); María José Paloma, MD (S. Hematología, H. Virgen del Camino, Pamplona, Spain); Nuria Bermejo, MD (S. Hematología, H. San Pedro de Alcántara, Cáceres, Spain); Rubén Berrueco, MD (S. Hematología, H. Sant Joan de Deu, Barcelona, Spain); José Mateo, MD (S. Hematología, H. Sta Creu i St Pau, Barcelona, Spain); Karmele Arribalzaga, MD (S. Hematología, H. Universitario Fundación Alcorcón, Madrid, Spain); Pascual Marco, MD (S. Hematología, H. General de Alicante, Alicante, Spain); Ángeles Palomo, MD (S. Hematología, H. Regional Universitario de Málaga, Málaga, Spain); Nerea Castro Quismondo, MD (S. Hematología, H. Universitario 12 de Octubre, Madrid, Spain); Belén Iñigo, MD (S. Hematología, H. Clínico San Carlos, Madrid, Spain); María del Mar Nieto, MD (S. Hematología, Complejo Hospitalario de Jaén, Jaén, Spain); Rosa Vidal, MD (S. Hematología, Fundación Jiménez Díaz, Madrid, Spain); María Paz Martínez, MD (S. Hematología, H. Nuestra Sra. de Sonsoles, Ávila, Spain); Reyes Aguinaco, MD (S.

Hematología, H. Joan XXIII, Tarragona, Spain); Maria Tenorio, MD (S. Hematología, H. Ramón y Cajal, Madrid, Spain); María Ferreiro, MD (S. Hematología, H. Montecelo, Pontevedra, Spain); Javier García-Frade, MD (S. Hematología, H. Río Hortega, Valladolid, Spain; Ana María Rodríguez-Huerta, MD (S. Hematología, H. Gregorio Marañón, Madrid, Spain); Jorge Cuesta, MD (S. Hematología, H. Virgen de la Salud, Toledo, Spain); Ramón Rodríguez-González, MD (S. Hematología, H.Severo Ochoa, Madrid, Spain); Faustino García-Candel, MD (S. Hematología, H. Universitario Virgen Arrixaca, Murcia, Spain); Manuela Dobón, MD (S. Hematología, H. Lozano Blesa, Zaragoza, Spain); Carlos Aguilar, MD (S. Hematología, H. Santa Bárbara, Soria, Spain); José María Bastida, MD (S. Hematología, H. Universitario, Salamanca, Spain); Ramiro Núñez, MD (S. Hematología, H. Universitario Virgen del Rocio, Sevilla, Spain); María Azanzazu Aguirre, MD (S. Hematología, H. Universitario Donostia, Spain); Marta Rodríguez-Luaces, MD (S. Hematología, H. Verge de la Cinta, Tortosa, Spain); Marina Gordillo, MD (S. Hematología, H. Doctor Jose Molina Orosa, Lanzarote, Spain); Carmen de Cos, MD (S. Hematología, H. Universitario Puerta del Mar, Cádiz, Spain); Laura Torres, MD (S. Hematología, Complejo Hospitalario Universitario Insular Materno-Infantil de Canarias, Las Palmas de Gran Canaria, Spain); Fernando Batlle (Department of Research,Lapisoft SA, A Coruña, Spain)

Discussió

“El objeto de toda discusión no debe ser el triunfo, sino el progreso.”

Joseph Antoine René Joubert

DISCUSSIÓ

La VWD és un trastorn de la coagulació causat per defectes hereditaris que afecten a la concentració, l'estructura o la funció del VWF. En aquesta tesi doctoral s'ha estudiat de manera exhaustiva la base genètica i la relació genotip i fenotip en el registre espanyol de la VWD, PCM-EVW-ES, i s'ha col·laborat en l'estudi genètic de la cohort portuguesa de pacients amb VWD. En base als resultats obtinguts i discutits extensament en els diferents manuscrits que conformen aquesta tesi, es pretén donar a continuació una visió global de les actuals dificultats diagnòstiques, esmentar les noves propostes i recomanacions per millorar la classificació dels diferents subtipus i discutir les futures perspectives en el complex terreny del diagnòstic d'aquesta malaltia.

1 Dificultats en el diagnòstic clínic de la VWD

Durant les últimes dècades l'aproximació diagnòstica de la VWD s'ha mantingut pràcticament invariable tot i les dificultats reportades en la identificació i la classificació dels pacients amb aquesta patologia, especialment degut a la heterogeneïtat de símptomes i la variabilitat intrínseca de les proves de laboratori emprades.[18, 217]

Un dels obstacles a l'hora d'establir un diagnòstic correcte consisteix en la valoració de la simptomatologia hemorràgica. En aquest sentit, la introducció, fa tretze anys, del primer qüestionari per avaluar el sagnat de forma objectiva s'ha considerat una eina molt útil per a la valoració dels pacients.[21] No obstant, donat que els símptomes hemorràgics més freqüents a la VWD, com epistaxis, hematomes i menorràgies, també s'acostumen a detectar en població sana, no es poden considerar com a un únic marcador pel diagnòstic d'aquesta patologia. Una bona mostra d'això és, per exemple, que la freqüència de menorràgia observada en dones diagnosticades de VWD, es similar a l'observada en poblacions de dones control.[18]

Addicionalment, degut a la manca d'una tècnica de referència de laboratori pel diagnòstic de la VWD, és necessari recórrer a la realització d'un conjunt extens de proves que suposen un considerable cost i temps dedicat al diagnòstic per a la classificació definitiva. Els tests principals inclouen la quantificació dels nivells d'antigen del VWF (VWF:Ag) i FVIII (FVIII:C) i la mesura de la funció del VWF, principalment mitjançant l'assaig VWF:RCo. Aquest últim test és poc reproduïble, presenta una elevada variabilitat intra- i inter-laboratoris i una baixa sensibilitat a nivells VWF:Ag <10 IU/dL que compromet la precisió del diagnòstic.[18, 217] Tanmateix, una altra complicació pròpia d'aquest assaig és la interferència amb el polimorfisme p.Asp1472His del *VWF*.[218] Aquest, present en dos terços de la població Afro-Americana i en una sisena part de la població caucàsica, provoca una disminució de la unió de la ristocetina al VWF, fet condueix a establir la conclusió errònia d'un VWF disfuncional, malgrat no s'associï amb l'observació d'una clínica de sagnat.[219] En base a aquestes troballes, considerem que l'aplicació de nous assajos que permeten avaluar la unió del VWF al receptor GPIb de les plaquetes sense la presència de ristocetina serà probablement el test d'elecció per a avaluar la funcionalitat del factor de forma més sensible i precisa en un futur pròxim.[220, 221] Aquests es basen en l'ús de recombinants de la proteïna GPIb amb mutacions de guany de funció i s'anomenen VWF:GPIbM.

Per altra banda, malgrat que el test VWF:Ag és més sensible que el VWF:RCo, la gran variabilitat dels nivells de VWF:Ag en la població general dificulten i converteixen en tot un repte la confirmació del diagnòstic de la VWD de tipus 1. Per exemple, l'augment gradual dels nivells del VWF amb l'edat (1-1,2% per any) complica la interpretació dels resultats, pel fet que en alguns casos poden arribar fins al rang de normalitat.[219] Tanmateix, l'elevada influencia del grup sanguini en els nivells del VWF condueix a la possibilitat que individus de la població general amb grup sanguini O presentin nivells de VWF:Ag just per sota del rang de normalitat (50 IU/dL) i, per tant, puguin ser diagnosticats erròniament de VWD.[37]

Finalment, és remarcable que moltes de les proves considerades de segon nivell, que permeten una classificació acurada de la VWD en subtipus, només estan disponibles en

hospitals de referència. A Espanya, malgrat la demostrada experiència dels centres en el diagnòstic d'aquesta patologia, els resultats d'una enquesta realitzada l'any 2011 posava de manifest la dificultat per accedir a algunes d'aquestes tècniques. Per exemple, la prova per analitzar els multímers estava posada a punt únicament en 5 hospitals (20%) i el test RIPA, indispensable pel diagnòstic de la VWD de tipus 2B, es realitzava sempre en 17 centres (50%) a concentracions de 1-1,5% mg/mL de ristocetina i en tan sols 10 centres (30%) a concentracions més baixes (0,2-0,5% mg/mL).[65] En aquest test és important fer un rang de concentracions per determinar el llindar de sensibilitat reactiva a la ristocetina. La hipersensibilitat a la ristocetina, provocat per l'augment de l'afinitat del VWF per unir-se al GPIb de les plaquetes, es indicatiu de tipus 2B VWD. No obstant, s'ha de tenir present que la realització del conjunt de tests de rutina de primera i segona línia no assegura la identificació de tots els pacients amb VWD. Aquest és el cas dels pacients amb mutacions que alteren la unió del VWF als col·làgens de tipus IV o VI, que passarien desapercebuts donat que la prova específica d'unió a col·lagen (VWF:CB) de tipus IV o VI està relegada a l'àmbit d'investigació.

Conjuntament, aquestes dificultats relatives a la clínica del pacient i a les proves de laboratori també s'han observat en els 38 centres hospitalaris espanyols que han participat en el registre PCM-EVW-ES. En concret, dels 556 pacients inclosos al registre, només 484 havien pogut ser classificats per tipus en els centres locals. Addicionalment, degut a les limitacions dels recursos disponibles, en 33 pacients diagnosticats a nivell local amb VWD de tipus 2 no s'havia especificat el subtipus (2A, 2B, 2M o 2N). Per altra banda, després de repetir les proves de laboratori en els centres de referència i re-analitzar les mostres, tan sols es va confirmar la patologia en 442 individus, i la classificació per tipus de VWD establerta pels centres de referència es diferenciava de la determinada a nivell local en un 42,3% dels casos. Tots aquests resultats posen de manifest les dificultats en la caracterització dels pacients de VWD a nivell espanyol, degut fonamentalment a la gran variabilitat dels resultats obtinguts en les proves de laboratori.

2 NGS: una realitat en el diagnòstic molecular de la VWD

En els últims anys, l'estratègia de seqüenciació més utilitzada pel diagnòstic molecular de la VWD ha estat la seqüenciació basada en el mètode descrit per *Sanger*.[89] Aquesta ha estat emprada tant en la confirmació del diagnòstic en pacients i estudis familiars com en la caracterització del genotip en grans cohorts de pacients. [142, 186, 216] Un dels obstacles associats a l'aplicació d'aquesta tecnologia és l'elevat cost que suposa la seqüenciació completa del *VWF* per identificar la mutació responsable. Per aquest motiu, molts laboratoris han optat per seqüenciar diferencialment determinats exons segons el fenotip clínic i de laboratori del pacient (veure apartat 4.2.2 de la introducció).[89, 132]

A Espanya, en el moment del naixement del projecte PCM-EVW-ES, l'any 2010, la seqüenciació tradicional del *VWF* es realitzava en quatre centres. En tres d'ells mitjançant una seqüenciació diferencial del *VWF*, i tan sols en el nostre laboratori es realitzava la seqüenciació completa i directa del gen gràcies al desenvolupament d'un procés ràpid i senzill per identificar les mutacions candidates.[65, 142] Això feia evident que molts centres hospitalaris espanyols no disposaven de les tècniques per abordar el diagnòstic molecular de la VWD. No obstant, malgrat tenir l'opció de derivar una mostra per fer l'estudi genètic als centres especialitzats, un 50% mai sol·licitaven aquest tipus d'estudi.[65] Aquesta informació posava de manifest que la major part de la població espanyola diagnosticada amb VWD no havia estat caracteritzada genèticament, fet que evidència la mancança del diagnòstic genètic assistencial a nivell espanyol.

Les tècniques per abordar el diagnòstic molecular de la VWD, a diferència dels test de laboratori, han anat evolucionat constantment i, des del 2005, s'ha produït un canvi substancial en les tècniques de seqüenciació del DNA amb la introducció de la NGS.[146] Aquesta tecnologia permet produir un gran volum d'informació de seqüència a un cost molt més baix que la seqüenciació de *Sanger*, oferint la possibilitat d'estudiar el *VWF* complet a un cost impensable fins al moment i assequible per molts laboratoris.[145, 147] Per tal d'aprofitar els avantatges d'aquesta tecnologia, ja des de l'any 2009 al nostre

laboratori es va començar a treballar en el desenvolupament de procediments basats en NGS per abordar el diagnòstic molecular de la VWD, especialment per l'estudi de grans cohorts de forma ràpida i econòmica.

En un primer estudi pilot per examinar la viabilitat de la seqüenciació del *VWF* amb la tecnologia NGS es van provar dues estratègies diferents: i) la seqüenciació completa del *VWF* per LR-PCR; i ii) la seqüenciació d'exons i regions intròniques flanquejants per PCR convencionals. Els resultats obtinguts van posar de manifest la dificultat per analitzar les variants intròniques profundes ja que la implicació clínica d'aquestes variants resultava de difícil interpretació.[142, 222] S'ha de tenir present que en el moment del desenvolupament d'aquest estudi, les bases de dades, com la de 1000G iniciada al 2008,[223] on es recopila informació dels variants i de la seva freqüència en diferents poblacions, estaven en desenvolupament i recollien poca informació, fet que dificultava la recerca de variants intròniques per verificar si aquestes havien estat detectades prèviament en individus sans, el que hauria ajudat a descartar o admetre el possible efecte patogènic.[89] Tanmateix, les eines de predicció *in silico* no tenen el poder per detrminar l'efecte deleteri de les variants intròniques.

En base a aquesta experiència prèvia del nostre laboratori, es va dissenyar el protocol basat en NGS presentat en aquest treball. Per aquest motiu, per a l'estudi molecular del *VWF* es van seleccionar els 52 exons del gen, així com, les regions intròniques flanquejants. Tanmateix, derivat dels resultats obtinguts en tres estudis genètics de la VWD de tipus 1 en grans cohorts on s'evidenciava que mutacions en la regió promotora del gen podrien ser responsables de la patologia, es va decidir incloure la regió promotora del *VWF* en el disseny.[185, 186, 202] De fet, actualment, gràcies a estudis *in vitro* i *in vivo* s'ha pogut demostrar que, com a mínim, una de les mutacions identificades en el promotor (c.-1522_-1510del13) causa una reducció de l'expressió del *VWF*.[158]

Un cop seleccionades les regions d'interès del *VWF* a ser analitzades, es va apostar per la innovadora tecnologia de microfluids (Access Array™ System (Fluidigm Corporation))[224] per abordar l'amplificació d'aquestes i la preparació de llibreries, necessària per a la posterior seqüenciació mitjançant NGS. La tecnologia de microfluids

integrava una sèrie de característiques favorables per a l'estudi del *VWF*, com per exemple, la capacitat per realitzar 48 PCRs per 48 pacients simultàniament. Addicionalment, la construcció de llibreries amb aquesta estratègia ha resultat senzilla i ràpida gràcies a un sistema de doble amplificació o PCR de fusió que permet l'addició simultània dels índexs, per discriminar els amplicons derivats de cada pacient, i dels adaptadors per poder seqüenciar les llibreries en una plataforma de NGS. Finalment, malgrat que aquest protocol requereix d'un equipament específic no disponible al nostre laboratori, no va ser un impediment per a la selecció d'aquesta tecnologia ja que al nostre entorn diferents hospitals i centres d'investigació disposaven de l'equipament necessari. La validació d'aquest protocol en mostres, prèviament estudiades per seqüenciació tradicional de *Sanger,* ha demostrat una sensibilitat i especificitat gairebé del 99% i 98%, respectivament, evidenciant l'eficiència del protocol dissenyat per a la seva aplicació al diagnòstic genètic del *VWF*.

Tot i que inicialment l'estudi molecular del *VWF* no estava contemplat en el registre espanyol de la VWD (PCM-EVW-ES) degut a l'elevat cost que suposava seqüenciar el gen mitjançant la tecnologia de *Sanger*, gràcies al desenvolupament del procediment presentat en aquest treball basat en NGS, s'ha pogut assumir el preu i s'ha realitzat el diagnòstic molecular en tots els pacients inclosos al registre PCM-EVW-ES. L'aplicació d'aquest protocol en l'estudi genètic de les cohorts espanyola i portuguesa de pacients amb VWD (articles 1-3) ha demostrat una sèrie d'avantatges evidents: i) reduir de manera significativa el material fungible i de reactius amb la conseqüent reducció de costos; ii) simplificar la creació de llibreries ja que es fa de forma simultània per a cada DNA durant el procés d'amplificació; i iii) evitar la necessitat d'una posterior normalització de les mostres.

Aquest és el primer procediment dissenyat per a l'anàlisi molecular del *VWF* aprofitant les avantatges de la tecnologia NGS. Des de la seva presentació al 2015 (article 1), en la literatura tan sols s'ha descrit una nova aproximació empleada per a la caracterització genètica d'una cohort xinesa de 90 pacients amb VWD.[225]

3 Utilitat dels estudis de grans cohorts en la VWD

Els estudis de grans cohorts són una de les eines més emprades i útils per aprofundir en el coneixement de la VWD, així com per tenir una imatge de l'epidemiologia molecular de la patologia en les diferents poblacions. El registre espanyol presentat en aquest treball va ser l'estudi prospectiu analitzat a nivell genètic més extens, incloent tots els tipus de VWD. L'estudi molecular realitzat en aquesta cohort per part del nostre laboratori posiciona la recerca espanyola en l'avantguarda del coneixement en aquest camp. Paral·lelament, la presentació de la cohort portuguesa amb 60 pacients no relacionats diagnosticats de VWD proporciona les primeres descripcions d'aquesta patologia a Portugal.

Un dels aspectes més importants a l'hora de realitzar estudis de grans cohorts és l'establiment d'uns bons criteris d'inclusió que permetran seleccionar de forma precisa la mostra d'individus a estudiar i establir el diagnòstic final en cada un d'ells. Fins al moment, en el cas de la VWD la majoria d'aquests han seguit els criteris establerts en les guies de la ISTH-SSC i NHLBI (*National Heart, Lung, and blood Institute*).[17, 18] No obstant, els criteris d'inclusió de la VWD de tipus 1 han variat lleugerament entre els diferents estudis publicats. En aquest sentit, actualment hi ha un debat obert en la comunitat científica sobre quin és el llindar dels nivells de VWF òptim per establir un diagnòstic definitiu pels pacients de tipus 1.[219] Per exemple, considerar com a veritables VWD de tipus 1 aquells que presenten nivells de VWF inferiors a 30 UI/dL ja que s'associen a mutacions al *VWF*, i considerar com a pacients amb factor de risc de sagnat aquells amb nivells baixos d'entre 30-50 UI/dL, donat que no s'acostuma a associar amb mutacions al *VWF*.[219] En aquest context, diverses cohorts han utilitzat nivells de VWF:Ag <50 IU/dL com a criteri d'inclusió,[185] i en altres estudis, com el registre francès, de VWF:Ag <30 IU/dL per a seleccionar els pacients que presenten sagnats més recurrents i excloure aquells que presenten una clínica lleu.[216] D'acord amb aquestes premisses, al registre PCM-EVW-ES s'ha seleccionat un llindar de VWF:Ag <30 IU/dL com un dels criteris d'inclusió, com s'especifica en l'article 1 i 2 d'aquesta tesi doctoral. Per altra banda, també s'ha defensat per part d'altres autors que l'algoritme diagnòstic

recomanat per confirmar la VWD, incloent la VWD de tipus 1, es basa en el BS, juntament amb un llindar <30 IU/dL pel test VWF:RCo i/o FVIII:C.[214, 226, 227] Aquests criteris d'inclusió s'han aplicat a la cohort portuguesa pel reclutament de pacients, tal i com s'indica en l'article 3.

Un altre dels punts a destacar dels estudis de grans cohorts, on participen varis centres d'una regió geogràfica, és la importància de la centralització de les mostres en els centres de referència, perseguint l'objectiu últim d'establir un diagnòstic homogeni. En aquest sentit, la centralització de les mostres de la cohort espanyola i portuguesa ha presentat avantatges significatius des del punt de vista de les proves de laboratori ja que han permès caracteritzar de forma més precisa a la població d'estudi evitant la variabilitat inter-laboratori descrita del test VWF:RCo. Per altra banda, des del punt de vista de l'anàlisi molecular del *VWF*, l'anàlisi simultània d'un gran número de mostres ha permès optimitzar la capacitat màxima dels seqüenciadors, el que s'ha traduït en un abaratiment del cost per mostra.[228] Per contra, aquelles proves que es realitzen *in situ* no s'han pogut reproduir en els centres de referència. Aquest es el cas del test RIPA que es realitza a partir del plasma fresc dels pacients i, per tant, només es pot fer en els centres hospitalaris en el moment de reclutar el pacient.[27]

Addicionalment, els resultats obtinguts en l'estudi d'ambdues cohorts ens ha permès tenir una imatge integral de la VWD, així com investigar la vàlua del l'estudi molecular del *VWF* per establir un diagnòstic precís en aquesta patologia. Durant molts anys i inclús a dia d'avui, varis autors qüestionen el potencial de l'estudi molecular del *VWF* a l'hora d'establir un correcte diagnòstic de la patologia per a tots el tipus de VWD i actualment encara no es considerada per alguns experts com una tècnica de rutina de primera línia. [27, 132, 229] De fet, alguns autors consideren que l'anàlisi genètic en la VWD només es justificada per la VWD de tipus 2, perquè pot influir en el tractament d'elecció, i de tipus 3, per proporcionar consell genètic i diagnòstic prenatal.[52, 217] Per contra, renuncien a realitzar l'estudi genètic en la VWD de tipus 1 ja que només s'identifiquen mutacions en una petita proporció de casos, i sol requerir l'anàlisi del *VWF* complet, que resulta molt costós mitjançant la seqüenciació tradicional.[52, 140, 217]

En aquest context, contràriament als arguments presentat, els resultats obtinguts en el projecte PCM-EVW-ES reforcen l'indiscutible paper de l'estudi genètic en la confirmació del diagnòstic i la classificació de la patologia en tots els tipus de VWD. La seva aplicació ha estat essencial per confirmar el diagnòstic d'aquesta coagulopatia en 480 dels 556 inicialment reclutats. En concret, ha permès la re-inclusió de 43 pacients descartats en base a l'anàlisi del fenotip en els centres de referència, la exclusió de 5 pacients on es sospitava de síndrome de VWD adquirida i la reclassificació de 105 pacients. A continuació s'exposen alguns exemples on l'estudi genètic ha tingut un paper decisiu. En primer lloc, ha permès classificar 35 pacients d'un total de 44 dels quals no es disposava del test RIPA, en 25 pacients de tipus 2A i 10 de tipus 2B, gràcies a la identificació de mutacions prèviament caracteritzades en la literatura. En segon lloc, l'anàlisi genètic ha resultat molt rellevant per diagnosticar correctament aquells pacients que presenten mutacions que afecten a la unió a col·lagen IV o VI degut a que els tests de laboratori VWF:CB4 o VWF:CB6 no es contemplen en l'algoritme de diagnòstic de rutina de la patologia.[171, 230] Tanmateix, s'ha de tenir en compte que en determinades ocasions aquests pacients poden no presentar alteracions en VWF:Ag i/o VWF:RCo i, per tant, podrien passar desapercebuts i sense diagnòstic. Aquest és el cas de 3 pacients de 7 pertanyents al PCM-EVW-ES amb mutacions d'alteració d'unió a col·lagen IV o VI que haurien quedat exclosos si ens haguéssim basat únicament amb els resultats de l'anàlisi del fenotip. Per consegüent, el nostres resultats semblen indicar que la freqüència de pacients que presenten la VWD causada per mutacions d'unió a col·lagen podria estar subestimada globalment. En tercer lloc, el diagnòstic genètic ha resultat essencial per la reclassificació acurada de determinats pacients, com és el cas de quatre individus inicialment diagnosticats com a tipus 1 i reclassificats com a tipus 2B després de la identificació de mutacions 2B Malmö. Aquestes mutacions són les p.Pro1266Leu/p.Pro1266Gln, i donat que no sempre s'associen amb una reducció dels HMWM, característica específica de la VWD de tipus 2B, la seva classificació resulta confusa si només es te en compte el fenotip observat.

Com a conseqüència derivada, en base al gran avantatge que suposa disposar d'un diagnòstic molecular del *VWF* juntament amb la reducció de costos de seqüenciació

que s'aconsegueix amb l'aplicació de la NGS, hem proposat un canvi de paradigma de l'algoritme de diagnòstic de la VWD amb la incorporació de l'estudi genètic del *VWF* com un test de primera línia, juntament amb l'avaluació dels sagnats, la mesura dels nivells de VWF i l'activitat funcional. En aquest sentit, la identificació de la mutació podria orientar sobre quines són les proves de segona línia que s'han de realitzar per confirmar el diagnòstic, sense la necessitat de recórrer a la execució de tot el conjunt de tests per caracteritzar el fenotip, el que suposaria una reducció tant en el temps com en els costos per establir la correcta classificació del pacient. Per tant, considerem que aquesta proposta de modificació de l'algoritme tindrà efectes molt beneficiosos tant en el diagnòstic, pronòstic i en el tractament de la VWD (veure apartat 4.1 de la introducció).

4 Avaluació dels resultats de l'estudi PCM-EVW-ES

L'aplicació del protocol dissenyat al registre PCM-EVW-ES ha permès realitzar l'estudi genètic en tots els pacients i conèixer la base genètica en cada un d'ells. Disposar d'una població ben caracteritzada a nivell genètic i fenotípic, ens ha permès establir la correlació entre les mutacions identificades i les característiques fenotípiques observades.

4.1 Anàlisi de les variants identificades al VWF

4.1.1 Mutacions al *VWF*

Al projecte PCM-EVW-ES s'han identificat un total de 704 variants (237 úniques), de les quals 155 no havien estat descrites prèviament a la base de dades internacional de la VWD (LOVD-VWD-EAHAD) (https://grenada.lumc.nl/LOVD2/VWF/variants.php). Aquests resultats suposen una valuosa informació per comprendre la base genètica d'aquesta patologia i ens han permès tenir una imatge més nítida dels mecanismes moleculars responsables de la VWD. S'ha demostrat que les mutacions causants de la VWD de tipus 1 i 2 acostumen a ser de tipus *missense*. Per contra, en la VWD de tipus

3 trobem un ventall més ampli de tipus de mutacions incloent: potencials mutacions d'*splicing* (PSSM), *nonsense*, *frameshift*, grans delecions i conversions gèniques. Aquest estudi corrobora l'elevada heterogeneïtat genètica associada a la VWD.

La identificació de dues grans delecions en el registre PCM-EVW-ES es especialment destacable ja que aquest tipus de mutacions són relativament infreqüents en la VWD. Aquestes es poden originar com a conseqüència de recombinacions homòlogues, a vegades involucrant seqüències repetitives Alu distribuïdes al llarg del gen, o no homòlogues.[133] La majoria de les grans delecions interrompen el marc de lectura, el que condueix a una pèrdua d'expressió de l'al·lel del *VWF* i s'associen habitualment amb VWD de tipus 3. Aquest és el cas de la mutació c.(1945+1_1946-1)_(7437+1_7438-1)del identificada en homozigosi en un pacient de tipus 3 del registre que provoca la deleció des de l'exó 16 al 43 del *VWF*. La identificació d'aquesta mutació suposa una aportació important a les bases moleculars de la VWD ja que, fins al moment, només s'han descrit una vintena de grans delecions associades a la VWD de tipus 3. [133, 231] Ara bé, les grans delecions també poden provocar la pèrdua d'aminoàcids sense alterar el marc de lectura (*inframe*), com la mutació p.Gly386_Ser1873del, identificada en un pacient de tipus 2 que provoca la deleció des de l'exó 11 fins al 32 com a conseqüència d'una recombinació homòloga entre seqüències de l'intró 10 i 32. Des del nostre coneixement, aquesta és la tercera gran deleció que s'associa a aquest tipus de VWD. Prèviament només dues havien estat descrites en la literatura: la deleció de l'exó 33 al 34 i la deleció de l'exó 26 al 34. [231, 232]

Una de les mutacions més característiques de la VWD, pel fet de no ser gaire comuns en altres patologies, són les conversions gèniques com a resultat de la recombinació entre el *VWF* (chr12) i el seu pseudogèn (chr22). Aquest tipus de mutació es veu afavorida per l'elevada homologia entre el gen i el pseudogèn i probablement, per la presència de seqüències anomenades *chi* que promouen recombinacions no homòlogues entre gens situats en cromosomes diferents.[233] En concret, en el *VWF* es coneixen dues seqüències *chi* localitzades a l'intró 27 i a l'exó 28.[234] Les conversions gèniques poden incloure des del canvi d'un sòl nucleòtid en la seqüència del *VWF* fins

a múltiples substitucions depenent de l'extensió reemplaçada pel pseudogèn.[164] En el registre PCM-EVW-ES s'ha identificat una gran conversió gènica que inclou cinc substitucions p.[(V1279I;Q1311*;I1343V;V1360A;F1369I)] en homozigosi en un pacient de tipus 3. Aquest tipus de conversió ja havia estat descrita prèviament en tres pacients no relacionats provinents de l'Índia.[164] Tanmateix, també s'han identificat substitucions puntuals com la p.Pro1266Leu en la VWD de tipus 2B i la p.Gln1311Ter en la VWD de tipus 3 que malgrat podrien ser degudes a canvis produïts a l'atzar, no es pot excloure la possibilitat que també es tracti de conversions gèniques.

En un 22% dels casos amb VWD de tipus 1 del projecte PCM-EVW-ES s'ha identificat la mutació p.Arg1205His que provoca un augment de l'aclariment del VWF de la circulació.[157] La identificació d'aquesta mutació és important des del punt de vista terapèutic ja que en aquests pacients el tractament amb desmopressina està contraindicat, donat que malgrat augmenti la concentració del VWF a la circulació, gràcies a l'alliberació del contingut de les cèl·lules endotelials, la proteïna serà eliminada ràpidament tornant als nivells basals en poc temps.[235]

Una conseqüència important derivada dels resultats obtinguts es que, en determinats casos, la identificació de la mutació ens pot fer qüestionar el tipus d'herència establerta segons el tipus de VWD. En aquest sentit, en el projecte PCM-EVW-ES i en la cohort portuguesa, s'ha observat que malgrat la VWD de tipus 3 es considerada d'herència autosòmica recessiva, determinats individus portadors de mutacions associades a aquest tipus de VWD han presentat simptomatologia hemorràgica i nivells reduïts de VWF. N'és un exemple el pacient del PCM-EVW-ES portador de la mutació p.Arg2434Ter, en heterozigosi, i diagnosticat de VWD de tipus 1, degut a la reducció dels nivells del VWF. Aquesta mutació havia estat prèviament descrita en homozigosi en pacients amb VWD de tipus 3 provinents de l'Índia,[236] motiu pel qual esperaríem que individus que presentessin la mutació en heterozigosi fossin diagnosticats com a portadors de mutacions de tipus 3. Recolzant-nos en aquests resultats i en d'altres reportats prèviament,[213] considerem que en determinats casos la VWD de tipus 3 no és clarament recessiva i s'hauria de considerar com a herència co-dominant .

Per altra banda, en algunes ocasions, la identificació de més d'una mutació en un pacient o la presència d'una única mutació en dominis multifuncionals del VWF causen efectes pleiotròpics i produeixen fenotips únics amb característiques de més d'un subtipus de VWD.[215, 237] Aquesta peculiar propietat restringida a casos puntuals ha estat identificada als pacients que presenten la mutació nova p.Arg763Ser, en heterozigosi, localitzada en el lloc d'escissió del propèptid. Aquests han estat classificats com a tipus 2A/2M en base al patró multimèric *smeary*, tot i que, també presenten una alteració en VWF:FVIIIB pròpia d'un tipus 2N.

La complementació de l'estudi mutacional amb l'anàlisi de lligament mitjançant amb marcadors STR ens ha permès traçar el possible origen comú de les mutacions recurrents al registre PCM-EVW-ES. [129, 133] La mutació p.Gln1311Ter ha estat una de les més freqüents identificada en 7 pacients de tipus 3, i en combinació amb p.Gly19Arg en un pacient de tipus 1. Els resultats de l'anàlisi de lligament revelen un haplotip comú en tots ells. En un estudi previ on es van analitzar 4 individus espanyols (2 famílies) d'origen Romaní que presentaven la mutació p.Gln1311Ter es va postular que aquesta te un origen fundador comú prèviament a la dispersió de la població gitana al nord-oest de l'Índia (any 1025).[238] Tanmateix, aquesta mutació ha estat identificada en 5/15 famílies amb VWD de tipus 3 pertanyents a la cohort de Portugal, de les quals almenys una és d'origen romaní. L'anàlisi dels marcadors genètics en tots els pacients portadors de la mutació p.Gln1311Ter conjuntament permetria una estimació més precisa de l'edat de la mutació. Addicionalment, l'ampliació de l'anàlisi de lligament a la cohort portuguesa de pacients esdevindria valuós ja que varies mutacions, a part de la p.Gln1311Ter, són comunes en ambdues cohorts, i degut a la proximitat geogràfica entre ambdós països, és molt probable que presentin un origen comú. Aquest és el cas de la mutació p.Cys2235ArgfsTer8, no descrita a la base de dades LOVD-VWD-EAHAD ni a la literatura, que ha estat identificada en un pacient del PCM-EVW-ES, i alhora en cinc pacients no relacionats de la cohort portuguesa. Per altra banda, la mutació de la primera descripció històrica de la VWD, c.2435delC (veure apartat 1 de la introducció), ha estat identificada en heterozigosi en un pacient del registre PCM-EVW-ES. Aquesta mutació és molt freqüent al nord d'Europa i els resultats d'un publicació de l'any 1993,

on es van realitzar anàlisis de lligament del marcador genètic situat al promotor, suggereixen un origen comú per aquesta mutació.[16] No obstant, el pacient del registre no és aparentment d'origen nòrdic, suggerint un origen independent de la mutació. Finalment, en el moment de la presentació d'aquesta tesi, s'està treballant amb les dades obtingudes de l'anàlisi de les quatre STRs en la cohort de pacients del PCM-EVW-ES amb l'objectiu de postular l'origen i datar l'aparició de determinades mutacions recurrents en base a l'estimació de la taxa de mutació dels microsatèl·lits i del temps de generació.

4.1.2 Els SNPs: grans desconeguts en la modulació de la VWD

La seqüenciació completa del *VWF* en pacients i familiars del PCM-EVW-ES ens ha proporcionat, informació no només de les mutacions responsables, sinó també informació genotípica de més de 100 polimorfismes en regió codificant del *VWF*. Tot això ens situa en una posició privilegiada per a la investigació d'una hipòtesi, plasmada en l'article 5, on es postula la participació dels polimorfismes (SNPs, de l'anglès *single nucleotide polimorphism*) genètics del *VWF* en la modulació de la gravetat i les manifestacions clíniques de la VWD. Estudis recents en individus sans proporcionen evidències de que els SNPs no sinònims poden contribuir de forma independent a la variabilitat fenotípica dels nivells de VWF i FVIII,[218, 239, 240] per la qual cosa és probable que aquests factors puguin contribuir a la penetració incompleta i la variabilitat de la presentació clínica de VWD, especialment de tipus 1. Per explorar aquesta idea en la cohort espanyola de pacients amb VWD, s'han seleccionat els 10 SNPs no sinònims més freqüents en aquesta població (MAF>5%) i s'ha investigat l'associació d'aquests amb diversos paràmetres del fenotip de la VWD. Els resultats obtinguts mostren una associació significativa entre els SNPs p.Thr789Ala, p.Val471Ile, p.Thr1381Ala i p.Gln852Arg i les mesures VWF:Ag, VWF:RCo i/o VWF:CB, que es calcula que representen des del 1,5% al 2,6% de la variància fenotípica de cada característica. Segons el nostre coneixement, encara que s'han realitzat estudis similars en la població general, aquesta és la major investigació realitzada fins al moment en una mostra de pacients amb VWD per abordar l'impacte dels SNPs del *VWF* en el fenotip d'aquesta

patologia. Els nostres resultats avalen els primers indicis de la implicació dels SNPs com a moduladors de la VWD i suposen un punt de partida per aprofundir en la investigació en aquest àmbit en població malalta. L'extensió d'aquest estudi a les famílies del projecte PCM-EVW-ES millorarà la potència estadística d'aquest anàlisi i permetran la identificació de nous SNP implicats en el fenotip de la VWD.

4.2 Estudi de la correlació genotip-fenotip al PCM-EVW-ES

Amb l'objectiu d'establir la correlació genotip-fenotip s'han contrastat els resultats de les proves de laboratori amb les mutacions identificades, tenint en compte la localització de les mutacions en el *VWF* i les descripcions prèvies en la literatura i la base de dades. Convé destacar que en aquest estudi s'ha establert una bona correlació genotip-fenotip en un 94,6% dels casos. En la major part de pacients de tipus 2 i tipus 3 es van identificar mutacions candidates que podrien explicar el seu fenotip. L'elevada correlació observada en la VWD de tipus 2, ha estat en part afavorida per la utilització de la classificació de la VWD de tipus 2A/2M. Considerem que la acceptació d'aquest subtipus per part de la ISTH ajudaria a la classificació dels pacients que presenten un patró de bandes normal amb una reducció relativa dels multímers d'alt pes molecular i que actualment són difícils de classificar ja que no compleixen les característiques fenotípiques estrictes per a ser classificats ni com a tipus 2A ni com a 2M. Aquest és el cas de les mutacions recurrents p.Arg1374His, p.Arg1374Cys i p.1315Cys que a la literatura es troben associades tant a tipus 2A com a tipus 2M.[59]

En 13 (8,2%) dels 159 pacients de tipus 1 no es va identificat mutació candidata. No obstant, es destacable l'elevada taxa de detecció de mutacions en comparació amb altres cohorts de tipus 1 VWD on no s'identifica mutació en aproximadament un 30% dels casos (Figura 19).[185, 186, 202] Aquesta diferència podria explicar-se pels criteris més restringents d'inclusió amb uns nivells de VWF:Ag <30 IU/dL, però també per la tecnologia de seqüenciació emprada.

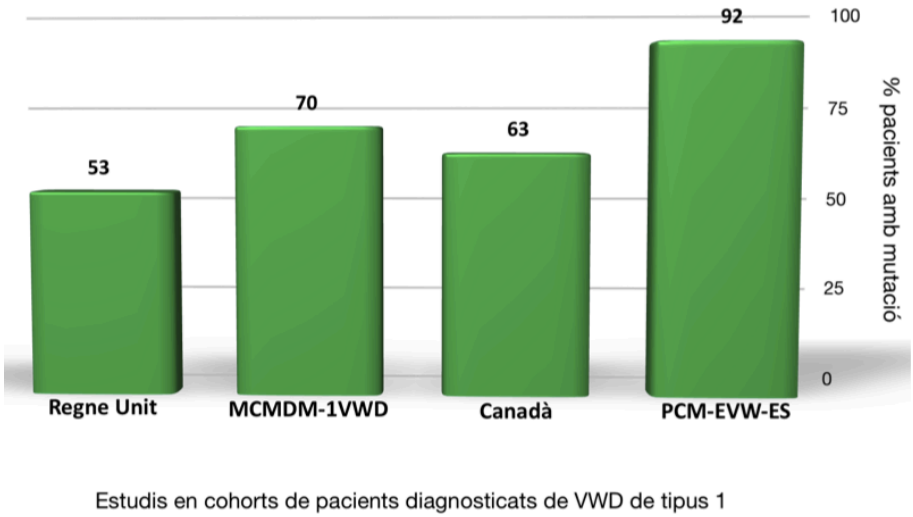


Figura 19. Taxa de detecció de variacions als estudis de VWD de tipus 1. *Percentatge dels pacients diagnosticats amb VWD de tipus 1 amb mutació candidata identificada en cadascun dels 4 grans estudis de VWD. La primera columna correspon a l'estudi del Regne Unit on s'analitzen pacients que compleixin el criteri d'inclusió VWF:RCo<50IU/dL. La segona columna representa l'estudi MCMDM-1VWD on s'analitzen pacients diagnosticats amb VWD de tipus 1 amb VWF:Ag<45 IU/dL. La tercera columna representa l'estudi canadenc de tipus 1 que inclou pacients amb VWF:Ag entre 5-50 IU/dL. La última columna fa referència als pacients de tipus 1 VWD inclosos a l'estudi PCM-EVW-ES presentat en aquesta Tesi Doctoral on s'han utilitzat uns criteris d'inclusió VWF:Ag<30.* [241]

La no identificació de mutacions candidates en el registre PCM-EVW-ES es podria atribuir a la limitació de la tècnica per identificar mutacions en regions intròniques profundes, o bé, a la presència de mutacions en altres *loci* no analitzats. En aquest sentit, la classificació de la VWD revisada al 2006 per la ISTH contempla la possibilitat que la VWD no estigui restringida a mutacions al *VWF*.[17] Aquesta hipòtesis es va plantejar ja a partir del 1999 en base als resultats publicats per varis estudis on demostraven que en algunes famílies les característiques fenotípiques de la VWD de tipus 1 no co-segregaven amb el *VWF*.[242-244] Un dels estudis més extensos, amb 143 famílies, remarca que en aproximadament el 50% dels pedigrís analitzats, la VWD de tipus 1 no co-segrega amb els marcadors genètics del *VWF,* assenyalant la possibilitat que altres gens diferents al *VWF* podrien determinar la reducció dels nivells de VWF.[245] En aquest sentit, a banda de la influencia àmpliament descrita del grup sanguini ABO en els nivells de VWF (veure apartat 2.1.2 de la introducció), estudis recents en població sana han identificat nous factors genètics implicats en la modulació dels nivells del VWF. En concret, estudis d'associació del genoma complet (GWAS, de les sigles en

anglès *Genome-Wide Association Study*) com el CHARGE han identificat els següents *loci*: *SCARA5*, *STAB2*, *TCN2* i *CLEC4M*.[246, 247] En base a aquests coneixements, estem treballant per reclutar el màxim número de membres familiars dels pacients del projecte PCM-EVW-ES sense mutació identificada per realitzar anàlisi de lligament. Els resultats d'estudiar els familiars afectes i sans, de com a mínim dues generacions, permetran identificar els pacients candidats de presentar alteracions en altres *loci*, implicats en el desenvolupament de la patologia. Actualment, aquest camp està molt poc explorat, no obstant, considerem que aquestes avaluacions en un futur no molt llunyà s'integraran en l'algoritme de diagnòstic d'aquesta patologia (veure apartat 9 de la discussió). Fet que representarà un avenç significatiu en la comprensió de les bases moleculars d'aquesta complexa patologia.

5 Importància de les bases de dades en el diagnòstic molecular de la VWD

L'evolució de les bases de dades de mutacions ha anat en paral·lel amb la de les tècniques de seqüenciació. Tot sembla indicar que el continuat desenvolupament de mètodes basats en NGS, com el presentat en aquest treball, produiran una major taxa de submissió de variants genètiques associades a malalties a les bases de dades. Tanmateix, aquesta tendència també es veurà afavorida pels estudis de seqüenciació a escala poblacional de pacients, com és el cas dels dos cohorts presentades en aquest treball.[248] Disposar de bases de dades cada vegada més completes és fonamental des del punt de vista diagnòstic, ja que la identificació de mutacions prèviament associades a la patologia representa una valuosa informació per confirmar la seva patogenicitat. Per aquest motiu, informar de les mutacions identificades, tant noves com reportades prèviament, a les bases de dades sempre ha estat recomanat encaridament a tots els investigadors.[249]

La base de dades EAHAD-VWD-LOVD ha estat un clar reflex de l'evolució que han anat tenint les bases de dades en els últims anys, el contingut de la qual ha variat

significativament tant en número com en tipus de mutacions incloses. L'any 2001 les mutacions de tipus 2 representaven el 70% del total de mutacions, i les de tipus 1 tan sols un 4,7% malgrat és el subtipus més freqüent a la població (60-80%). Aquesta diferencia s'explica pel fet que molts laboratoris optaven per examinar regions limitades del *VWF* degut a l'elevat cost que suposava la seqüenciació completa del *VWF* mitjançant *Sanger*, sent els exons 28 i del 18-20 els més analitzats.[133] Deu anys més tard, al 2011, gràcies als resultats obtinguts en els tres grans estudis multicèntrics de la VWD de tipus 1 realitzats la passada dècada, es va augmentar significativament el número de mutacions de tipus 1, les quals representaven llavors el 28,8%.[185, 186, 202, 249] En l'actualitat, hi ha un total de 1193 registres, dels quals 708 corresponen a mutacions úniques. Si revisem l'aportació dels laboratoris espanyols a la base de dades, prèvia al projecte PCM-EVW-ES, observem que tan sols representen un 4,5% del total d'entrades (53 mutacions totals), sent la majoria enviades per part del nostre laboratori (37 mutacions). No obstant, recentment el total de 704 mutacions identificades al PCM-EVW-ES han estat enviades a la base de dades EAHAD-VWD-LOVD, de manera que representaran un 38% del total de registres, així com, a la base de dades Human Genome Mutation Database (HGMD), on es recopiles mutacions associades a diferents malalties genètiques.[248] L'aportació del registre PCM-EVW-ES en ambdues bases de dades, ens situa a les posicions capdavanteres de contribució al coneixement de la etiologia molecular d'aquesta patologia.

També juguen un paper rellevant al diagnòstic molecular les bases de dades que recopilen informació genòmica de població sana, on s'indica la freqüència al·lèlica de les variants en diverses poblacions mundials, com la base de dades de 1.000 Genomes o *l'Exome Aggregation Consortium*. Registres de aquest tipus han sofert un creixement exponencial com a conseqüència de l'arribada de la NGS, ja que permet la seqüenciació d'exomes i genomes a un cost molt reduït. Aquestes bases de dades, cada vegada més completes, proporcionen una informació sense precedents per a la comunitat científica ja que permeten distingir els canvis patògens de les variants benignes senzillament en base a la seva freqüència en les diferents poblacions.[250] En aquest sentit, determinades variants *missense* del *VWF* (p.His817Gln, p.Met740Ile i p.Arg2185Gln) inicialment

considerades mutacions patogèniques en pacients europeus, es van identificar posteriorment en Afro-Americans sans amb una freqüència al·lèlica entre 10-15%, fet que indicaria que no són responsables de la VWD.[240]

Finalment, per a l'administració de registres i projectes multicèntrics amb informació clínica, genètica i de laboratori és imprescindible comptar amb eines col·laboraries preferiblement en línia que permetin la interacció dels investigadors implicats. En aquest sentit i amb l'objectiu de recopilar sistematitzadament la informació dels pacients inclosos al registre PCM-EVW-ES, englobant les dades clíniques i fenotípiques recollides en els centres locals, els estudis de coagulació realitzats en els centres de referència i la informació genètica, s'ha creat una base de dades d'accés online (https://www.proyectopcm.com/) amb una estructura a mida d'acord amb els requeriments específics del projecte . Aquesta base de dades és d'accés restringit limitat als hematòlegs i investigadors dels hospitals que han participat a l'estudi. No obstant, s'està treballant en l'opció de que, una part de la informació de la base de dades esdevingui d'accés públic ja que proporciona dades molt rellevants sobre l'epidemiologia de la VWD a Espanya que poden suposar un suport addicional al diagnòstic.

6 Investigant l'efecte patogènic de les mutacions

En tota anàlisi genètica un cop identificada una variant candidata s'ha de determinar si aquesta és responsable del desenvolupament de la patologia. Per desxifrar l'efecte de les mutacions s'utilitzen diferents aproximacions incloent els estudis *in silico*, estudis transcripcionals de l'mRNA i/o estudis funcionals d'expressió.

6.1 Estudis *in silico*

Aquests estudis s'han considerat molt útils en el diagnòstic molecular de les malalties monogèniques ja que són una primera aproximació ràpida i econòmica per determinar el possible efecte patogènic de les mutacions. Tanmateix, degut a l'augment de

l'aplicació de la NGS en l'àmbit clínic i el conseqüent increment de detecció de mutacions noves, l'ús d'aquests algoritmes de predicció s'ha considerat molt adequat com a suport a la priorització de variants candidates, per determinar quines tenen una major probabilitat de ser deletèries i que requeriran un seguiment addicional.[251] En base a aquestes premisses, totes les mutacions *missense*, potencials mutacions d'*splicing* (PSSM) i sinònimes identificades en ambdues cohorts de pacients amb VWD d'aquest treball s'han analitzat *in silico* amb un mínim de tres programes. La comparació de l'efecte deleteri conegut de mutacions caracteritzades prèviament en la literatura amb els resultats obtinguts amb l'anàlisi *in silico*, han demostrat que la predicció de les mutacions *missense* és més precisa que l'estimada per les PSSM. Aquest és el cas de la predicció de l'efecte de la mutació sinònima c.546G>A (article 3), on dos dels programes utilitzats han predit un efecte en l'*splicing*, mentre que els altres dos no. Tot i això, creiem que l'evolució dels anàlisis de predicció *in silico*, en base a una millor comprensió de les seqüències involucrades en el processament d'*splicing* juntament amb una millora de la integritat i l'organització de les dades de variants patogèniques conegudes associades a malaltia, farà que aquestes eines siguin més exactes i indispensables per establir un diagnòstic molecular precís. No obstant, actualment encara no són suficients i és requereixen d'estudis funcionals per confirmar l'efecte patogènic real de les mutacions.

6.2 Estudis transcripcionals del processament de l'mRNA

Aquests estudis són especialment útils per determinar l'efecte de les potencials mutacions d'*splicing* sobre els transcrits dels gens candidats. Les PSSM són totes aquelles que poden alterar el normal processament del pre-mRNA. Aquestes inclouen: les substitucions produïdes en les regions consensus d'*splicing*, que afecten als dinucleòtids GT i AG dels extrems 5' i 3' de cada intró; les intròniques i; les mutacions que recauen en regió exònica, com les de tipus *missense* o sinònimes.[252] En aquest sentit, s'ha descrit que al voltant d'un 15% de les mutacions puntuals que recauen en exons provoquen alteracions en l'*splicing* del mRNA donant lloc al desenvolupament de malalties genètiques humanes. [253]

En el cas de la VWD, s'han reportat menys de 30 PSSM a la base de dades de la VWD (https://grenada.lumc.nl/LOVD2/VWF/variants.php), la majoria afectant a la regió consens d'*splicing*. No obstant, pocs laboratoris són els que han investigat el seu efecte *in vivo* o *in vitro* sobre el mRNA del *VWF*. En aquest context, els resultats presentats en l'article 4 conjuntament amb els publicats prèviament pel nostre laboratori[191] representen la contribució més extensa als estudis funcionals *in vivo* mitjançant mRNA amb un total de 23 PSSM localitzades al *VWF*. En concret, en l'article 4, s'han caracteritzat quatre mutacions localitzades en la regió consens d'*splicing* amb un efecte clarament patogènic (c.3379+1G>A, c.1533+1G>A, c.5664+2T>C, c.7082-2A>G). Tanmateix, es va demostrar l'efecte de dues mutacions sinònimes, c.546G>A i c.7437G>A, en l'*splicing* del mRNA. Finalment, s'ha descrit la que és , fins al moment, la segona mutació *missense* que afecta al processament del mRNA del *VWF*, la mutació p.Cys370Tyr. Aquesta, localitzada a l'exó 9, promou la generació de dos transcrits aberrants: sense l'exó 9 i sense l'exó 8 i 9. Els nostres resultats representen, sota el nostre punt de vista, una interesant aportació als coneixements sobre els mecanismes moleculars que condueixen a la VWD i posen de manifest la necessitat d'estudiar determinades mutacions sinònimes i *missense* també a nivell de l'mRNA.

Per altra banda, en el nostre estudi, totes les mutacions que afecten a l'*splicing* s'han identificat en pacients diagnosticats amb la VWD quantitativa (tipus 1 o tipus 3). Tenint en compte que la majoria d'aquestes mutacions produeixen canvis en la pauta de lectura i la conseqüent introducció de codons de parada prematurs (PTCs), tot sembla indicar que el procés cel·lular anomenat *nonsense mediated decay* (NMD) juga un paper important en el mecanisme patogènic, conduint a la deficiència quantitativa del VWF. No obstant, és important recordar que en la literatura, tot i que en una proporció menor, s'han descrit mutacions que afecten al processament de l'mRNA en pacients diagnosticats amb VWD de tipus qualitatiu (tipus 2). Aquest és el cas, per exemple, de la mutació c.3538G>A, identificada en un pacient de tipus 2A, que dona lloc a la deleció *in-frame* de dos exons que codifiquen per part del domini D3, implicat en la formació dels multímers.[254]

Fins fa relativament pocs anys, els estudis per analitzar l'efecte de les mutacions del *VWF* s'acostumaven a realitzar en RNA de plaquetes ja que aquest és un dels tipus cel·lulars on s'expressa el *VWF* juntament amb les cèl·lules endotelials. No obstant, estudis anteriors realitzats en el nostre laboratori van posar de manifest la importància d'analitzar l'efecte de les mutacions en l'mRNA de leucòcits, on el VWF s'expressa de forma ectòpica, ja que oferia resultats complementaris.[191] Per aquest motiu, l'efecte de totes les mutacions estudiades en l'article 4 s'han analitzat en ambdós tipus cel·lulars, confirmant la utilitat d'aquesta aproximació.

Per últim, en base a l'experiència prèvia del laboratori en l'aplicació de la tecnologia NGS per a l'estudi de PSSM identificades al *F8*, es va optimitzar un protocol per abordar l'anàlisi de les PSSM al *VWF* utilitzant la NGS.[255] El nou procediment, que combina RT-PCR i el posterior anàlisis mitjançant NGS, ha demostrat ser més ràpid i més sensible que el mètode desenvolupat prèviament al nostre laboratori.[191] En base a aquests resultats, postulem que la NGS presentarà un paper indiscutible no tan sols en el diagnòstic molecular, sinó també en la caracterització de mutacions a nivell transcripcional.

6.3 Les BOECs com a model cel·lular per l'estudi de mutacions del *VWF*

Les perspectives en els estudis funcionals *in vitro* per a determinar l'efecte patogènic de mutacions del *VWF* han canviat substancialment durant els últims anys gràcies a la possibilitat de disposar cèl·lules endotelials obtingudes directament del pacient, les BOECs (de les sigles en anglès, *Blood Outgrowth Endothelial Cells*).[200] Donat que permeten investigar i caracteritzar l'efecte de mutacions del *VWF* al seu lloc natural d'expressió,[198, 201, 256] un altre dels objectius plantejats al projecte PCM-EVW-ES ha estat estudiar l'efecte patogènic de mutacions seleccionades, principalment de tipus *missense*, utilitzant les BOECs com a model cel·lular. Actualment ja s'han seleccionat 11 mutacions d'interès i, en el moment de finalitzar aquesta memòria, s'han aïllat BOECs de quatre individus control, dos pacients de tipus 1 i dos pacients de tipus 2. Cal

destacar que, a banda de la utilitat d'aquest model cel·lular per estudiar mutacions a nivell proteic, també és adequat per a investigar l'efecte de les PSSM ja que les BOECs representen una nova font per obtenir mRNA del *VWF* del pacient. Aquesta alternativa ha estat emprada satisfactòriament per a la caracterització de tres PSSM.[257] Tot sembla indicar que l'anàlisi concomitant de transcrits derivats de plaquetes, leucòcits i BOECs del pacient, juntament amb l'estudi funcional de la proteïna en BOECs, representaran un nou enfocament per a l'estudi integral del mecanisme patològic de les mutacions.

7 Altres abordatges pel diagnòstic molecular de la VWD: els panells de gens

Tot i que el procediment presentat en aquest treball no ha perdut vigència i continua sent una tècnica competitiva per abordar l'estudi del *VWF* en una gran cohort de pacients diagnosticats amb una mateixa patologia, en l'actualitat es disposa d'altres aproximacions per abordar l'estudi molecular de les malalties monogèniques com la VWD mitjançant NGS. [258] Aquest és el cas dels panells de gens que ens permeten plantejar l'estudi simultani de desenes o milers de gens a un cost molt inferior al de l'estudi individual de cada un d'ells mitjançant la tecnologia tradicional. El disseny d'aquests panells també es poden abordar, entre altres, mitjançant la tecnologia de Fluidigim que permet amplificar un elevat número de gens a un preu molt reduït *(*https://www.fluidigm.com/*).* Els panells de gens que inclouen l'anàlisi del *VWF* conjuntament amb altres gens de la cascada de la coagulació (o relacionats amb l'hemostàsia) presenten avantatges significatius des del punt de vista assistencial d'un laboratori de diagnòstic molecular, on es reben mostres de diferents coagulopaties congènites. En concret, aquesta aproximació ja ha estat explorada pel nostre laboratori gràcies al desenvolupament d'un panell de 23 gens, inclòs el *VWF*, per abordar el diagnòstic genètic integral de les coagulopaties congènites.[259] La implementació d'aquest protocol ha resultat molt versàtil i ha permès aplicar una mateixa tècnica per l'estudi de diferents patologies, simplificant protocols i reduint els costos de seqüenciació. Tanmateix, ha resultat especialment útil des del punt de vista del

diagnòstic de la VWD, pel fet que ha permès diferenciar de forma simple, ràpida i directa amb un únic protocol la VWD de les seves genocòpies. Aquest és el cas de la VWD de tipus 2N que presenta mutacions al *VWF* i dona lloc a una simptomatologia i nivells de factor idèntics a l'hemofília A (mutacions al *F8*) i el de la VWD de tipus 2B degut a alteracions al *VWF* envers la VWD de tipus plaquetari (PT-VWD) (mutacions al *GP1BA*). De fet, tots els resultats obtinguts amb el panell de gens (resultats no mostrats en aquesta tesi) demostren que l'abordatge de l'estudi molecular de patologies relacionades des del punt de vista clínic, genètic i/o bioquímic hauria d'avançar cap a l'anàlisi mitjançant solucions integrals. Per altra banda, els panells de gens també poden incorporar altres *loci* implicats en la modulació dels nivells de VWF, com els gens *STXBP5* i *STX2*, relacionats amb el tràfic vesicular i l'exocitosi, així com, *SCARA5*, *STAB2* i *CLEC4M*, que codifiquen per a receptors d'aclariment. Aquesta possibilitat resulta prometedora pel diagnòstic de la VWD en aquells pacients on no s'ha identificat mutació al *VWF*.[260] Actualment aquest camp està molt poc explorat en població malalta, no obstant, es podria esperar que la inclusió d'aquests gens al panell permetrien identificar mutacions i/o polimorfismes que podrien donar resposta i explicar el fenotip observat en el 30% dels pacients amb VWD de tipus 1 sense mutacions causals identificades al *VWF*.[185, 186, 202, 261]

8 Perspectives de futur

El projecte PCM-EVW-ES està tenint una gran transcendència a Espanya i internacional ja que és la primera cohort on s'ha aplicat un protocol basat en NGS per realitzar el diagnòstic molecular. A més, la creació d'aquest registre espanyol ha permès caracteritzar de forma exhaustiva (genètica i fenotípicament) 556 pacients diagnosticats de VWD, donant suport a les mancances en el diagnòstic d'aquesta patologia a Espanya. Motivat per l'èxit del projecte i gràcies a la consecució de nou finançament, s'ha reobert el registre per incloure al voltant de 400 nous pacients. En el moment de la presentació d'aquest treball, s'estan reclutant i obtenint les mostres dels pacients inclosos en aquesta segona fase del registre. En els pròxims mesos es durà a

terme l'estudi acurat del fenotip, així com del genotip. Per l'abordatge de l'estudi molecular del VWF es valorarà l'opció d'utilitzar el procediment presentat en aquest treball, o bé, del disseny d'un panell de gens.

Per altra banda, en l'apartat de perspectives de futur, no podem deixar d'esmentar la evolució que està seguint la tecnologia i l'impacte que això tindrà en els pròxims anys. El diagnòstic molecular de les malalties amb base genètica està canviant radicalment, a una velocitat difícilment assimilable per l'esser humà des de la irrupció de les tecnologies de seqüenciació massiva. En aquest sentit, quan encara no ens hem familiaritzat amb el que suposa el diagnòstic mitjançant la seqüenciació de varis gens simultàniament amb els panells de gens, altres aproximacions com la seqüenciació de l'exoma clínic (CES), on s'analitzen gens relacionats amb patologia humana, o l'exoma complet (WES), on es seqüencia tota la regió codificant de proteïnes del genoma, són eines cada vegada més emprades pel diagnòstic molecular de malalties d'herència mendeliana poc freqüents. Encara més, probablement en els pròxims anys el mètode de diagnòstic de rutina per a totes les patologies amb base genètica, com la VWD, es basarà en la seqüenciació del genoma complet (WGS). El cost del WGS ha anat caient molt ràpidament des de la seqüenciació del primer genoma. Actualment el WGS amb alguns aparells de la plataforma Illumina ja està per sota dels 1.000 dòlars.[145] Propiciat per la disponibilitat de la seqüenciació del genoma humà a un cost assequible ja s'han iniciat varis projectes nacionals que inclouen la seqüenciació de genomes, com per exemple: *100.000 Genomes Project of Genomics England* (Regne Unit), *100.000 Genomes Project* (Austràlia) i *Precision Medicine Initiative* (EE.UU). Tot i això, l'evolució dels seqüenciadors de segona generació, com el NovaSeq (Illumina) i, el desenvolupament de la tecnologia de seqüenciació de tercera generació permetran que la WGS per menys de 100 dòlars sigui una realitat durant la pròxima dècada. Aquesta esperada reducció del cost de seqüenciació farà que es plantegi com una opció real l'aplicació del WGS de manera rutinària com un test en els recent nascuts en un futur no massa llunyà.[262, 263]

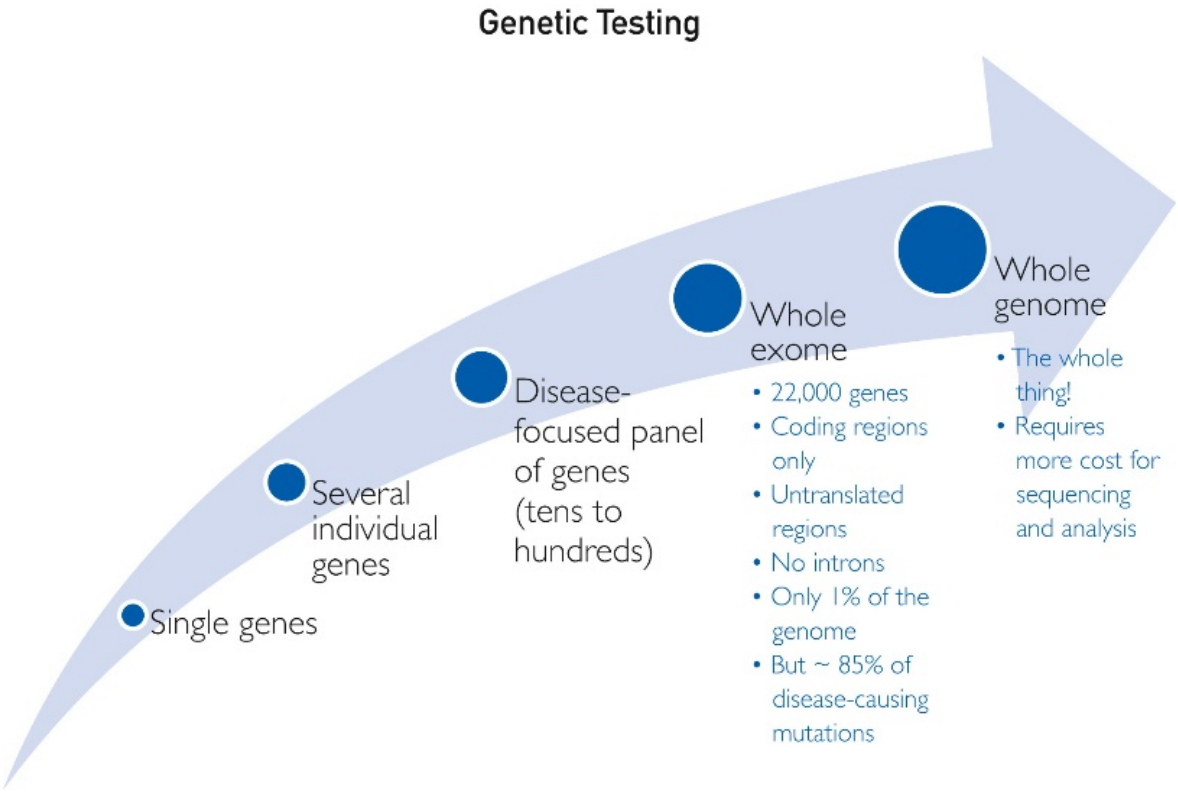
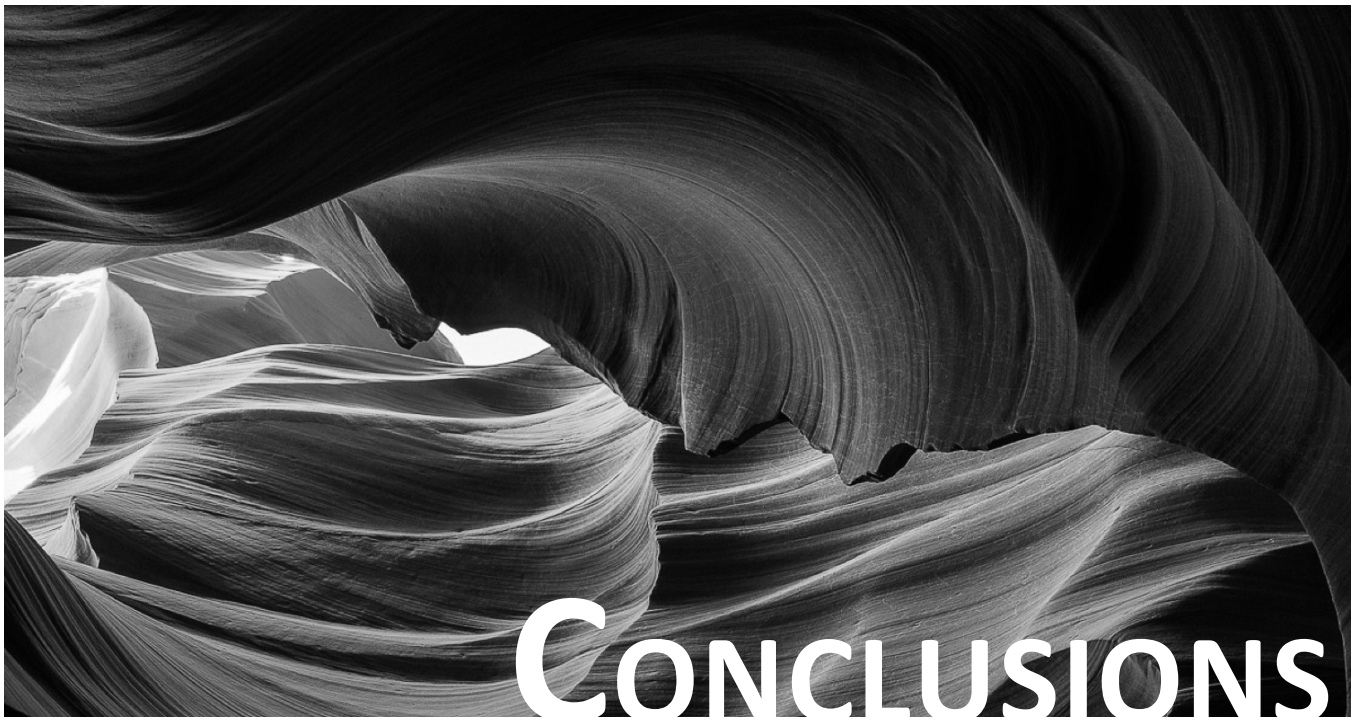


Figura 20. Evolució del diagnòstic genètic des del 1990 a l'actualitat. *Imatge de Lindor et al.*[236]

Tots aquests antecedents, incloent els avenços en les tecnologies de seqüenciació així com la dramàtica reducció dels costos, avalen que el diagnòstic molecular en els pròxims anys passa per la implementació de la WGS en la sanitat pública, el que es traduirà en una medicina personalitzada. Si totes aquestes previsions es confirmen, la seqüenciació del *VWF* (o dels gens responsables de qualsevol altra malaltia monogènica) per identificar la mutació causal de la VWD no serà necessària, i es recorrerà a revisar la informació del WGS realitzada durant els primers dies de vida. Cal destacar que pel cas de la VWD, a banda de la identificació de mutacions i variacions al *VWF*, l'aplicació d'aquesta aproximació permetrà l'anàlisi dels gens modificadors discutits anteriorment i la detecció de nous *loci* responsables del sagnat en els pacients sense mutació en el *VWF*, fet que suposaria una millora en l'atenció dels pacients amb trastorns hemorràgics hereditaris.[264] En resum, és inqüestionable que en els propers anys les tecnologies de seqüenciació massiva continuaran tenint un efecte transformador en el diagnòstic genètic de les malalties amb base genètica com la VWD.



CONCLUSIONS

"El caràcter no pot ser desenvolupat en tranquil·litat i quietud. Només a través de l'experiència, les proves i el sofriment, l'ambició pot ser inspirada i l'exit assolit."

Helen Keller

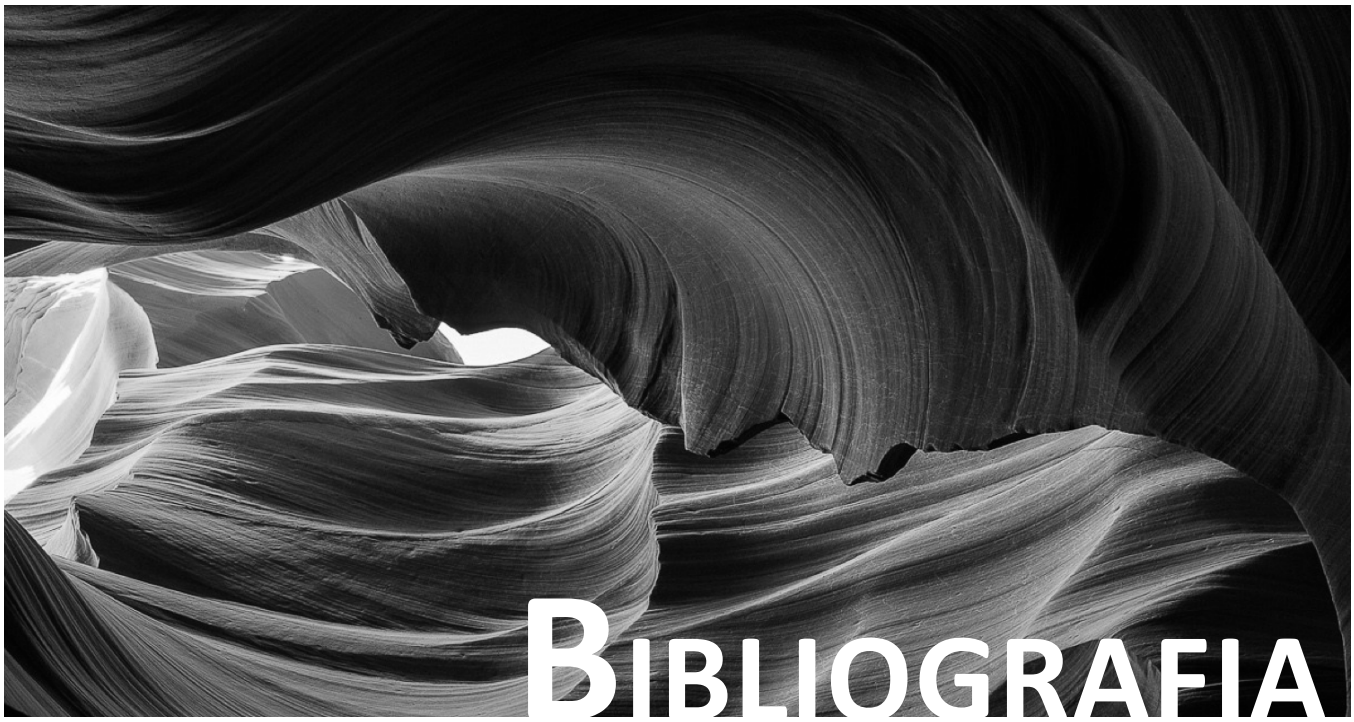
CONCLUSIONS

1. S'ha dissenyat un protocol per realitzar l'estudi molecular de la VWD que permet l'amplificació simultània del *VWF* de 48 pacients mitjançant la tecnologia Access Array™ de Fluidigm i la posterior seqüenciació en paral·lel de fins a 192 pacients en la plataforma MiSeq d'Illumina.
2. S'ha validat l'eficàcia i fiabilitat del protocol dissenyat mitjançant la identificació de totes les mutacions prèviament caracteritzades per *Sanger* en 172 pacients procedents de l'Hospital Universitari Vall d'Hebron i el Complex Hospitalari Universitari d'A Coruña.
3. L'aplicació del procediment al registre de pacients del PCM-EVW-ES ha permès la identificació d'un total de 237 mutacions diferents al *VWF*, de les quals 155 no havien estat descrites prèviament a la base de dades internacional de mutacions. La seqüenciació del *VWF* juntament amb els resultats de laboratori han permès confirmar el diagnòstic de la VWD en 480 pacients dels 556 inicialment inclosos, establint-se la correlació genotip-fenotip en un 94,6% dels pacients.
4. L'aplicació del procediment a la cohort portuguesa ha permès identificar totes les mutacions detectades prèviament per *Sanger* i 26 mutacions addicionals, demostrant la seva potència i eficàcia.en el diagnòstic molecular dels pacients.
5. Derivat dels dos punts anteriors podem concloure que l'anàlisi genètica del *VWF* mitjançant la tecnologia NGS ha demostrat la seva vàlua en el diagnòstic de la VWD en ambdues cohorts, el que ha permès plantejar un nou algoritme per a la identificació de pacients amb VWD on l'estudi molecular del *VWF* formi part dels test de primera línia de diagnòstic, juntament amb la història personal i familiar de sagnats, els nivells d'antigen (VWF:Ag) i els nivells d'activitat del VWF (VWF:RCo).

6. L'ús de les eines d'anàlisi *in silico* han esdevingut de gran utilitat a l'hora d'avaluar de forma ràpida i econòmica el potencial efecte deleteri de les mutacions identificades en ambdues cohorts. La seva aplicació ha estat especialment rellevant per predir l'efecte de les mutacions noves, així com per seleccionar les més interessants per realitzar estudis funcionals.

7. Els estudis transcripcionals *in vivo* realitzats per avaluar 18 potencials mutacions d'*splicing* han permès determinar l'efecte patogènic de 8 d'elles. L'aplicació de la NGS en aquest estudi ha estat essencial per a la identificació de trànscrits aberrants poc expressats. Els resultats obtinguts posen de manifest que les mutacions de tipus *missense* i sinònimes també poden produir el seu efecte patogènic alterant el processament de l'mRNA.

8. L'estudi dels polimorfismes del *VWF* en el projecte PCM-EVW-ES ha evidenciat l'associació de quatre polimorfismes (p.Thr789Ala, p.Val471Ile, p.Thr1381Ala i p.Gln852Arg) amb variacions en les mesures quantitatives relacionades amb la VWD, incloent els nivells d'antigen de VWF (VWF:Ag), d'activitat (VWF:RCo) i d'unió a col·lagen (VWF:CB).



BIBLIOGRAFIA

"Ets amo de les paraules que no has dit, però esclau de les que han sortit dels teus llavis"

Anònim

BIBLIOGRAFIA

1. Von Willebrand EA. Hereditary pseudohaemophilia. . Finska Lakarsallskapets. 1926;67(7-112.
2. Lassila R, Lindberg O. Erik von Willebrand. Haemophilia : the official journal of the World Federation of Hemophilia. 2013;19(5):643-647.
3. Berntorp E. Erik von Willebrand. Thrombosis research. 2007;120 Suppl 1(S3-4.
4. Berntorp E, Ingerslev J, Schulman S. Von Willebrand disease: an update in the Aland islands. Haemophilia : the official journal of the World Federation of Hemophilia. 1999;5 Suppl 2(1-6.
5. Nilsson IM. The history of von Willebrand disease. Haemophilia : the official journal of the World Federation of Hemophilia. 1999;5 Suppl 2(7-11.
6. Nilsson IM, Blomback M, Von Francken I. On an inherited autosomal hemorrhagic diathesis with antihemophilic globulin (AHG) deficiency and prolonged bleeding time. Acta medica Scandinavica. 1957;159(1):35-57.
7. Nilsson IM, Blomback M, Jorpes E, Blomback B, Johansson SA. Von Willebrand's disease and its correction with human plasma fraction 1-0. Acta medica Scandinavica. 1957;159(3):179-188.
8. Zimmerman TS, Ratnoff OD, Powell AE. Immunologic differentiation of classic hemophilia (factor 8 deficiency) and von Willebrand's dissase, with observations on combined deficiencies of antihemophilic factor and proaccelerin (factor V) and on an acquired circulating anticoagulant against antihemophilic factor. The Journal of clinical investigation. 1971;50(1):244-254.
9. Gralnick HR, Coller BS. Studies of the human factor VIII/von Willebrand's factor protein. II. Identification and characterization of the von Willebrand protein. Blood. 1975;46(3):417-430.
10. Weiss HJ, Rogers J, Brand H. Defective ristocetin-induced platelet aggregation in von Willebrand's disease and its correction by factor VIII. The Journal of clinical investigation. 1973;52(11):2697-2707.
11. Lynch DC, Zimmerman TS, Collins CJ, et al. Molecular cloning of cDNA for human von Willebrand factor: authentication by a new method. Cell. 1985;41(1):49-56.
12. Sadler JE, Shelton-Inloes BB, Sorace JM, Harlan JM, Titani K, Davie EW. Cloning and characterization of two cDNAs coding for human von Willebrand factor. Proceedings of the National Academy of Sciences of the United States of America. 1985;82(19):6394-6398.

13. Ginsburg D, Handin RI, Bonthron DT, et al. Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) clones and chromosomal localization. Science. 1985;228(4706):1401-1406.

14. Verweij CL, Diergaarde PJ, Hart M, Pannekoek H. Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protein considerably larger than the mature vWF subunit. The EMBO journal. 1986;5(8):1839-1847.

15. Blomback M. Scientific visits to the Aland Islands. Haemophilia : the official journal of the World Federation of Hemophilia. 1999;5 Suppl 2(12-18.

16. Zhang ZP, Blomback M, Nyman D, Anvret M. Mutations of von Willebrand factor gene in families with von Willebrand disease in the Aland Islands. Proceedings of the National Academy of Sciences of the United States of America. 1993;90(17):7937-7940.

17. Sadler JE, Budde U, Eikenboom JC, et al. Update on the pathophysiology and classification of von Willebrand disease: a report of the Subcommittee on von Willebrand Factor. Journal of thrombosis and haemostasis : JTH. 2006;4(10):2103-2114.

18. Nichols WL, Hultin MB, James AH, et al. von Willebrand disease (VWD): evidence-based diagnosis and management guidelines, the National Heart, Lung, and Blood Institute (NHLBI) Expert Panel report (USA). Haemophilia : the official journal of the World Federation of Hemophilia. 2008;14(2):171-232.

19. van Galen KP, Mauser-Bunschoten EP, Leebeek FW. Hemophilic arthropathy in patients with von Willebrand disease. Blood reviews. 2012;26(6):261-266.

20. Sanders YV, Fijnvandraat K, Boender J, et al. Bleeding spectrum in children with moderate or severe von Willebrand disease: Relevance of pediatric-specific bleeding. American journal of hematology. 2015;90(12):1142-1148.

21. Rodeghiero F, Castaman G, Tosetto A, et al. The discriminant power of bleeding history for the diagnosis of type 1 von Willebrand disease: an international, multicenter study. Journal of thrombosis and haemostasis : JTH. 2005;3(12):2619-2626.

22. Tosetto A, Rodeghiero F, Castaman G, et al. A quantitative analysis of bleeding symptoms in type 1 von Willebrand disease: results from a multicenter European study (MCMDM-1 VWD). Journal of thrombosis and haemostasis : JTH. 2006;4(4):766-773.

23. Bowman M, Mundell G, Grabell J, et al. Generation and validation of the Condensed MCMDM-1VWD Bleeding Questionnaire for von Willebrand disease. Journal of thrombosis and haemostasis : JTH. 2008;6(12):2062-2066.

24. Bowman M, Riddel J, Rand ML, Tosetto A, Silva M, James PD. Evaluation of the diagnostic utility for von Willebrand disease of a pediatric bleeding questionnaire. Journal of thrombosis and haemostasis : JTH. 2009;7(8):1418-1421.

25. Rodeghiero F, Tosetto A, Abshire T, et al. ISTH/SSC bleeding assessment tool: a standardized questionnaire and a proposal for a new bleeding score for inherited

bleeding disorders. Journal of thrombosis and haemostasis : JTH. 2010;8(9):2063-2065.

26. Elbatarny M, Mollah S, Grabell J, et al. Normal range of bleeding scores for the ISTH-BAT: adult and pediatric data from the merging project. Haemophilia : the official journal of the World Federation of Hemophilia. 2014;20(6):831-835.
27. Favaloro EJ, Pasalic L, Curnow J. Laboratory tests used to help diagnose von Willebrand disease: an update. Pathology. 2016;48(4):303-318.
28. Ingerslev J. A sensitive ELISA for von Willebrand factor (vWf:Ag). Scandinavian journal of clinical and laboratory investigation. 1987;47(2):143-149.
29. Veyradier A, Fressinaud E, Sigaud M, Wolf M, Meyer D. A new automated method for von Willebrand factor antigen measurement using latex particles. Thrombosis and haemostasis. 1999;81(2):320-321.
30. James PD, Goodeve AC. von Willebrand disease. Genetics in medicine : official journal of the American College of Medical Genetics. 2011;13(5):365-376.
31. Stockschlaeder M, Schneppenheim R, Budde U. Update on von Willebrand factor multimers: focus on high-molecular-weight multimers and their role in hemostasis. Blood coagulation & fibrinolysis : an international journal in haemostasis and thrombosis. 2014;25(3):206-216.
32. Sadler JE. A revised classification of von Willebrand disease. For the Subcommittee on von Willebrand Factor of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. Thrombosis and haemostasis. 1994;71(4):520-525.
33. Favaloro EJ. Utility of the von Willebrand factor collagen binding assay in the diagnosis of von Willebrand disease. American journal of hematology. 2017;92(1):114-118.
34. Sadler JE, Mannucci PM, Berntorp E, et al. Impact, diagnosis and treatment of von Willebrand disease. Thrombosis and haemostasis. 2000;84(2):160-174.
35. Harrison RL, McKee PA. Estrogen stimulates von Willebrand factor production by cultured endothelial cells. Blood. 1984;63(3):657-664.
36. Mahieu B, Jacobs N, Mahieu S, et al. Haemostatic changes and acquired activated protein C resistance in normal pregnancy. Blood coagulation & fibrinolysis : an international journal in haemostasis and thrombosis. 2007;18(7):685-688.
37. Gill JC, Endres-Brooks J, Bauer PJ, Marks WJ, Jr., Montgomery RR. The effect of ABO blood group on the diagnosis of von Willebrand disease. Blood. 1987;69(6):1691-1695.
38. Franchini M, Crestani S, Frattini F, Sissa C, Bonfanti C. ABO blood group and von Willebrand factor: biological implications. Clinical chemistry and laboratory medicine. 2014;52(9):1273-1276.
39. Jenkins PV, O'Donnell JS. ABO blood group determines plasma von Willebrand factor levels: a biologic function after all? Transfusion. 2006;46(10):1836-1844.

40. Sztukowska M, Gallinaro L, Cattini MG, et al. Von Willebrand factor propeptide makes it easy to identify the shorter Von Willebrand factor survival in patients with type 1 and type Vicenza von Willebrand disease. British journal of haematology. 2008;143(1):107-114.

41. Ng C, Motto DG, Di Paola J. Diagnostic approach to von Willebrand disease. Blood. 2015;125(13):2029-2037.

42. Mannucci PM, Federici AB. Antibodies to von Willebrand factor in von Willebrand disease. Advances in experimental medicine and biology. 1995;386(87-92.

43. Iorio A, Oliovecchio E, Morfini M, Mannucci PM, Association of Italian Hemophilia Centres D. Italian Registry of Haemophilia and Allied Disorders. Objectives, methodology and data analysis. Haemophilia : the official journal of the World Federation of Hemophilia. 2008;14(3):444-453.

44. James PD, Lillicrap D, Mannucci PM. Alloantibodies in von Willebrand disease. Blood. 2013;122(5):636-640.

45. Zimmerman TS, Dent JA, Ruggeri ZM, Nannini LH. Subunit composition of plasma von Willebrand factor. Cleavage is present in normal individuals, increased in IIA and IIB von Willebrand disease, but minimal in variants with aberrant structure of individual oligomers (types IIC, IID, and IIE). The Journal of clinical investigation. 1986;77(3):947-951.

46. Mazurier C, Dieval J, Jorieux S, Delobel J, Goudemand M. A new von Willebrand factor (vWF) defect in a patient with factor VIII (FVIII) deficiency but with normal levels and multimeric patterns of both plasma and platelet vWF. Characterization of abnormal vWF/FVIII interaction. Blood. 1990;75(1):20-26.

47. Schneppenheim R, Budde U, Krey S, et al. Results of a screening for von Willebrand disease type 2N in patients with suspected haemophilia A or von Willebrand disease type 1. Thrombosis and haemostasis. 1996;76(4):598-602.

48. Federici AB, Rand JH, Bucciarelli P, et al. Acquired von Willebrand syndrome: data from an international registry. Thrombosis and haemostasis. 2000;84(2):345-349.

49. Franchini M, Lippi G. Acquired von Willebrand syndrome: an update. American journal of hematology. 2007;82(5):368-375.

50. Federici AB, Budde U, Castaman G, Rand JH, Tiede A. Current diagnostic and therapeutic approaches to patients with acquired von Willebrand syndrome: a 2013 update. Seminars in thrombosis and hemostasis. 2013;39(2):191-201.

51. Sanders YV, Giezenaar MA, Laros-van Gorkom BA, et al. von Willebrand disease and aging: an evolving phenotype. Journal of thrombosis and haemostasis : JTH. 2014;12(7):1066-1075.

52. Flood VH, Christopherson PA, Gill JC, et al. Clinical and laboratory variability in a cohort of patients diagnosed with type 1 VWD in the United States. Blood. 2016;127(20):2481-2488.

53. Batlle J, Perez-Rodriguez A, Franqueira MD, Lopez-Fernandez MF. Type 2M von Willebrand disease: a variant of type 2A? Journal of thrombosis and haemostasis : JTH. 2008;6(2):388-390.

54. Hilbert L, Gaucher C, Mazurier C. Identification of two mutations (Arg611Cys and Arg611His) in the A1 loop of von Willebrand factor (vWF) responsible for type 2 von Willebrand disease with decreased platelet-dependent function of vWF. Blood. 1995;86(3):1010-1018.

55. Casonato A, Pontara E, Sartorello F, Bertomoro A, Durante C, Girolami A. Type 2M von Willebrand disease variant characterized by abnormal von willebrand factor multimerization. The Journal of laboratory and clinical medicine. 2001;137(1):70-76.

56. Penas N, Perez-Rodriguez A, Torea JH, et al. von Willebrand disease R1374C: type 2A or 2M? A challenge to the revised classification. High frequency in the northwest of Spain (Galicia). American journal of hematology. 2005;80(3):188-196.

57. Michiels JJ, Berneman Z, Gadisseur A, et al. Classification and characterization of hereditary types 2A, 2B, 2C, 2D, 2E, 2M, 2N, and 2U (unclassifiable) von Willebrand disease. Clinical and applied thrombosis/hemostasis : official journal of the International Academy of Clinical and Applied Thrombosis/Hemostasis. 2006;12(4):397-420.

58. Doruelo AL, Haberichter SL, Christopherson PA, et al. Clinical and laboratory phenotype variability in type 2M von Willebrand disease. Journal of thrombosis and haemostasis : JTH. 2017;15(8):1559-1566.

59. Gadisseur A, van der Planken M, Schroyens W, Berneman Z, Michiels JJ. Dominant von Willebrand disease type 2M and 2U are variable expressions of one distinct disease entity caused by loss-of-function mutations in the A1 domain of the von Willebrand factor gene. Acta haematologica. 2009;121(2-3):145-153.

60. Rodeghiero F, Castaman G, Dini E. Epidemiological investigation of the prevalence of von Willebrand's disease. Blood. 1987;69(2):454-459.

61. Werner EJ, Broxson EH, Tucker EL, Giroux DS, Shults J, Abshire TC. Prevalence of von Willebrand disease in children: a multiethnic study. The Journal of pediatrics. 1993;123(6):893-898.

62. Castaman G, Linari S. Diagnosis and Treatment of von Willebrand Disease and Rare Bleeding Disorders. Journal of clinical medicine. 2017;6(4):

63. Berliner SA, Seligsohn U, Zivelin A, Zwang E, Sofferman G. A relatively high frequency of severe (type III) von Willebrand's disease in Israel. British journal of haematology. 1986;62(3):535-543.

64. Srivastava A, Rodeghiero F. Epidemiology of von Willebrand disease in developing countries. Seminars in thrombosis and hemostasis. 2005;31(5):569-576.

65. Batlle J, Perez-Rodriguez A, Pinto JC, Fraga EL, Rodriguez-Trillo Tch A, Fernanda Lopez-Fernandez M. Diagnosis and management of von Willebrand disease in Spain. Seminars in thrombosis and hemostasis. 2011;37(5):503-510.

66. Kujovich JL. von Willebrand's disease and menorrhagia: prevalence, diagnosis, and management. American journal of hematology. 2005;79(3):220-228.

67. Pasi KJ, Collins PW, Keeling DM, et al. Management of von Willebrand disease: a guideline from the UK Haemophilia Centre Doctors' Organization. Haemophilia : the official journal of the World Federation of Hemophilia. 2004;10(3):218-231.

68. Mannucci PM. Treatment of von Willebrand's Disease. The New England journal of medicine. 2004;351(7):683-694.

69. Mannucci PM. Desmopressin (DDAVP) in the treatment of bleeding disorders: the first 20 years. Blood. 1997;90(7):2515-2521.

70. Kaufmann JE, Vischer UM. Cellular mechanisms of the hemostatic effects of desmopressin (DDAVP). Journal of thrombosis and haemostasis : JTH. 2003;1(4):682-689.

71. Montgomery RR, Gill JC. Interactions between von Willebrand factor and Factor VIII: where did they first meet. Journal of pediatric hematology/oncology. 2000;22(3):269-275.

72. Mazurier C, Gaucher C, Jorieux S, Goudemand M. Biological effect of desmopressin in eight patients with type 2N ('Normandy') von Willebrand disease. Collaborative Group. British journal of haematology. 1994;88(4):849-854.

73. Michiels JJ, van de Velde A, van Vliet HH, van der Planken M, Schroyens W, Berneman Z. Response of von Willebrand factor parameters to desmopressin in patients with type 1 and type 2 congenital von Willebrand disease: diagnostic and therapeutic implications. Seminars in thrombosis and hemostasis. 2002;28(2):111-132.

74. Holmberg L, Nilsson IM, Borge L, Gunnarsson M, Sjorin E. Platelet aggregation induced by 1-desamino-8-D-arginine vasopressin (DDAVP) in Type IIB von Willebrand's disease. The New England journal of medicine. 1983;309(14):816-821.

75. Revel-Vilk S, Schmugge M, Carcao MD, Blanchette P, Rand ML, Blanchette VS. Desmopressin (DDAVP) responsiveness in children with von Willebrand disease. Journal of pediatric hematology/oncology. 2003;25(11):874-879.

76. Smith TJ, Gill JC, Ambruso DR, Hathaway WE. Hyponatremia and seizures in young children given DDAVP. American journal of hematology. 1989;31(3):199-202.

77. Mannucci PM. How I treat patients with von Willebrand disease. Blood. 2001;97(7):1915-1919.

78. Berntorp E, Archey W, Auerswald G, et al. A systematic overview of the first pasteurised VWF/FVIII medicinal product, Haemate P/ Humate -P: history and

clinical performance. European journal of haematology Supplementum. 2008;70):3-35.

79. Harper P, Favaloro EJ, Curtin J, Barnes C, Dunkley S. Human plasma-derived FVIII/VWD concentrate (Biostate): a review of experimental and clinical pharmacokinetic, efficacy and safety data. Drugs in context. 2016;5(212292.

80. Franchini M, Mannucci PM. Von Willebrand factor (Vonvendi(R)): the first recombinant product licensed for the treatment of von Willebrand disease. Expert review of hematology. 2016;9(9):825-830.

81. Castillo R, Monteagudo J, Escolar G, Ordinas A, Magallon M, Martin Villar J. Hemostatic effect of normal platelet transfusion in severe von Willebrand disease patients. Blood. 1991;77(9):1901-1905.

82. Castillo R, Escolar G, Monteagudo J, Aznar-Salatti J, Reverter JC, Ordinas A. Hemostasis in patients with severe von Willebrand disease improves after normal platelet transfusion and normalizes with further correction of the plasma defect. Transfusion. 1997;37(8):785-790.

83. Mannucci PM, Kempton C, Millar C, et al. Pharmacokinetics and safety of a novel recombinant human von Willebrand factor manufactured with a plasma-free method: a prospective clinical trial. Blood. 2013;122(5):648-657.

84. Gill JC, Castaman G, Windyga J, et al. Hemostatic efficacy, safety, and pharmacokinetics of a recombinant von Willebrand factor in severe von Willebrand disease. Blood. 2015;126(17):2038-2046.

85. De Meyer SF, Deckmyn H, Vanhoorelbeke K. von Willebrand factor to the rescue. Blood. 2009;113(21):5049-5057.

86. De Meyer SF, Vanhoorelbeke K, Chuah MK, et al. Phenotypic correction of von Willebrand disease type 3 blood-derived endothelial cells with lentiviral vectors expressing von Willebrand factor. Blood. 2006;107(12):4728-4736.

87. Pergolizzi RG, Jin G, Chan D, et al. Correction of a murine model of von Willebrand disease by gene transfer. Blood. 2006;108(3):862-869.

88. Mancuso DJ, Tuley EA, Westfield LA, et al. Human von Willebrand factor gene and pseudogene: structural analysis and differentiation by polymerase chain reaction. Biochemistry. 1991;30(1):253-269.

89. Baronciani L, Goodeve A, Peyvandi F. Molecular diagnosis of von Willebrand disease. Haemophilia : the official journal of the World Federation of Hemophilia. 2017;23(2):188-197.

90. Hernandez-Zamora E, Zavala-Hernandez C, Quintana-Gonzalez S, Reyes-Maldonado E. [Von Willebrand disease. Molecular biology and diagnosis]. Cirugia y cirujanos. 2015;83(3):255-264.

91. Mancuso DJ, Tuley EA, Westfield LA, et al. Structure of the gene for human von Willebrand factor. The Journal of biological chemistry. 1989;264(33):19514-19527.

92. Sadler JE. Biochemistry and genetics of von Willebrand factor. Annual review of biochemistry. 1998;67(395-424.

93. Shelton-Inloes BB, Titani K, Sadler JE. cDNA sequences for human von Willebrand factor reveal five types of repeated domains and five possible protein sequence polymorphisms. Biochemistry. 1986;25(11):3164-3171.

94. Zhou YF, Eng ET, Zhu J, Lu C, Walz T, Springer TA. Sequence and structure relationships within von Willebrand factor. Blood. 2012;120(2):449-458.

95. Lenting PJ, Casari C, Christophe OD, Denis CV. von Willebrand factor: the old, the new and the unknown. Journal of thrombosis and haemostasis : JTH. 2012;10(12):2428-2437.

96. Rauch A, Wohner N, Christophe OD, Denis CV, Susen S, Lenting PJ. On the versatility of von Willebrand factor. Mediterranean journal of hematology and infectious diseases. 2013;5(1):e2013046.

97. Mazzucato M, Spessotto P, Masotti A, et al. Identification of domains responsible for von Willebrand factor type VI collagen interaction mediating platelet adhesion under high flow. The Journal of biological chemistry. 1999;274(5):3033-3041.

98. Springer TA. von Willebrand factor, Jedi knight of the bloodstream. Blood. 2014;124(9):1412-1425.

99. Romijn RA, Bouma B, Wuyster W, et al. Identification of the collagen-binding site of the von Willebrand factor A3-domain. The Journal of biological chemistry. 2001;276(13):9985-9991.

100. Leebeek FW, Eikenboom JC. Von Willebrand's Disease. The New England journal of medicine. 2016;375(21):2067-2080.

101. Lenting PJ, Christophe OD, Denis CV. von Willebrand factor biosynthesis, secretion, and clearance: connecting the far ends. Blood. 2015;125(13):2019-2028.

102. Swystun LL, Rydz N, Notley C, et al. Genetic variability of the CLEC4M endothelial vectin receptor modulates binding and internalization of von Willebrand Factor and contributes to variance in plasma VWF levels. Blood. 2012;120(

103. Lenting PJ, Westein E, Terraube V, et al. An experimental model to study the in vivo survival of von Willebrand factor. Basic aspects and application to the R1205H mutation. The Journal of biological chemistry. 2004;279(13):12102-12109.

104. Casari C, Lenting PJ, Wohner N, Christophe OD, Denis CV. Clearance of von Willebrand factor. Journal of thrombosis and haemostasis : JTH. 2013;11 Suppl 1(202-211.

105. Fujimura Y, Titani K, Holland LZ, et al. A heparin-binding domain of human von Willebrand factor. Characterization and localization to a tryptic fragment extending from amino acid residue Val-449 to Lys-728. The Journal of biological chemistry. 1987;262(4):1734-1739.

106. Ruggeri ZM. Structure of von Willebrand factor and its function in platelet adhesion and thrombus formation. Best practice & research Clinical haematology. 2001;14(2):257-279.

107. Peyvandi F, Garagiola I, Baronciani L. Role of von Willebrand factor in the haemostasis. Blood transfusion = Trasfusione del sangue. 2011;9 Suppl 2(s3-8.

108. Macfarlane RG. An Enzyme Cascade in the Blood Clotting Mechanism, and Its Function as a Biochemical Amplifier. Nature. 1964;202(498-499.

109. Davie EW, Ratnoff OD. Waterfall Sequence for Intrinsic Blood Clotting. Science. 1964;145(3638):1310-1312.

110. Monroe DM, Roberts HR, Hoffman M. Platelet procoagulant complex assembly in a tissue factor-initiated system. British journal of haematology. 1994;88(2):364-371.

111. Hoffman M, Monroe DM, 3rd. A cell-based model of hemostasis. Thrombosis and haemostasis. 2001;85(6):958-965.

112. Vlot AJ, Koppelman SJ, Bouma BN, Sixma JJ. Factor VIII and von Willebrand factor. Thrombosis and haemostasis. 1998;79(3):456-465.

113. Pober JS, Sessa WC. Evolving functions of endothelial cells in inflammation. Nature reviews Immunology. 2007;7(10):803-815.

114. Kawecki C, Lenting PJ, Denis CV. von Willebrand factor and inflammation. Journal of thrombosis and haemostasis : JTH. 2017;15(7):1285-1294.

115. Pendu R, Terraube V, Christophe OD, et al. P-selectin glycoprotein ligand 1 and beta2-integrins cooperate in the adhesion of leukocytes to von Willebrand factor. Blood. 2006;108(12):3746-3752.

116. Denis CV, Andre P, Saffaripour S, Wagner DD. Defect in regulated secretion of P-selectin affects leukocyte recruitment in von Willebrand factor-deficient mice. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(7):4072-4077.

117. Starke RD, Ferraro F, Paschalaki KE, et al. Endothelial von Willebrand factor regulates angiogenesis. Blood. 2011;117(3):1071-1080.

118. Saint-Lu N, Oortwijn BD, Pegon JN, et al. Identification of galectin-1 and galectin-3 as novel partners for von Willebrand factor. Arteriosclerosis, thrombosis, and vascular biology. 2012;32(4):894-901.

119. Fressinaud E, Meyer D. International survey of patients with von Willebrand disease and angiodysplasia. Thrombosis and haemostasis. 1993;70(3):546.

120. Bosmans JM, Kockx MM, Vrints CJ, Bult H, De Meyer GR, Herman AG. Fibrin(ogen) and von Willebrand factor deposition are associated with intimal thickening after balloon angioplasty of the rabbit carotid artery. Arteriosclerosis, thrombosis, and vascular biology. 1997;17(4):634-645.

121. Qin F, Impeduglia T, Schaffer P, Dardik H. Overexpression of von Willebrand factor is an independent risk factor for pathogenesis of intimal hyperplasia: preliminary studies. Journal of vascular surgery. 2003;37(2):433-439.

122. Li S, Wang Z, Liao Y, et al. The glycoprotein Ibalpha-von Willebrand factor interaction induces platelet apoptosis. Journal of thrombosis and haemostasis : JTH. 2010;8(2):341-350.

123. Mochizuki S, Soejima K, Shimoda M, et al. Effect of ADAM28 on carcinoma cell metastasis by cleavage of von Willebrand factor. Journal of the National Cancer Institute. 2012;104(12):906-922.

124. Bryckaert M, Rosa JP, Denis CV, Lenting PJ. Of von Willebrand factor and platelets. Cellular and molecular life sciences : CMLS. 2015;72(2):307-326.

125. Rao ES, Ng CJ. Current approaches to diagnostic testing in von Willebrand Disease. Transfusion and apheresis science : official journal of the World Apheresis Association : official journal of the European Society for Haemapheresis. 2018;

126. Othman M, Kaur H, Emsley J. Platelet-type von Willebrand disease: new insights into the molecular pathophysiology of a unique platelet defect. Seminars in thrombosis and hemostasis. 2013;39(6):663-673.

127. James P, Lillicrap D. The role of molecular genetics in diagnosing von Willebrand disease. Seminars in thrombosis and hemostasis. 2008;34(6):502-508.

128. Shelton-Inloes BB, Chehab FF, Mannucci PM, Federici AB, Sadler JE. Gene deletions correlate with the development of alloantibodies in von Willebrand disease. The Journal of clinical investigation. 1987;79(5):1459-1465.

129. Vidal F, Julia A, Altisent C, Puig L, Gallardo D. Von Willebrand gene tracking by single-tube automated fluorescent analysis of four short tandem repeat polymorphisms. Thrombosis and haemostasis. 2005;93(5):976-981.

130. Peake IR, Bowen D, Bignell P, et al. Family studies and prenatal diagnosis in severe von Willebrand disease by polymerase chain reaction amplification of a variable number tandem repeat region of the von Willebrand factor gene. Blood. 1990;76(3):555-561.

131. van Amstel HK, Reitsma PH. Tetranucleotide repeat polymorphism in the vWF gene. Nucleic acids research. 1990;18(16):4957.

132. Keeney S, Bowen D, Cumming A, et al. The molecular analysis of von Willebrand disease: a guideline from the UK Haemophilia Centre Doctors' Organisation Haemophilia Genetics Laboratory Network. Haemophilia : the official journal of the World Federation of Hemophilia. 2008;14(5):1099-1111.

133. Goodeve AC. The genetic basis of von Willebrand disease. Blood reviews. 2010;24(3):123-134.

134. Keeney S, Cumming A, Hay C. Mutations in von Willebrand factor multimerization domains are not a common cause of classical type 1 von Willebrand disease. Thrombosis and haemostasis. 1999;82(5):1446-1450.

135. Hashemi Soteh M, Peake IR, Marsden L, et al. Mutational analysis of the von Willebrand factor gene in type 1 von Willebrand disease using conformation sensitive gel electrophoresis: a comparison of fluorescent and manual techniques. Haematologica. 2007;92(4):550-553.

136. Kakela JK, Friedman KD, Haberichter SL, et al. Genetic mutations in von Willebrand disease identified by DHPLC and DNA sequence analysis. Molecular genetics and metabolism. 2006;87(3):262-271.

137. Ribba AS, Lavergne JM, Bahnak BR, Derlon A, Pietu G, Meyer D. Duplication of a methionine within the glycoprotein Ib binding domain of von Willebrand factor detected by denaturing gradient gel electrophoresis in a patient with type IIB von Willebrand disease. Blood. 1991;78(7):1738-1743.

138. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America. 1977;74(12):5463-5467.

139. Kircher M, Kelso J. High-throughput DNA sequencing--concepts and limitations. BioEssays : news and reviews in molecular, cellular and developmental biology. 2010;32(6):524-536.

140. Peyvandi F, Kunicki T, Lillicrap D. Genetic sequence analysis of inherited bleeding diseases. Blood. 2013;122(20):3423-3431.

141. James P, Lillicrap D. Genetic testing for von Willebrand disease: the Canadian experience. Seminars in thrombosis and hemostasis. 2006;32(5):546-552.

142. Corrales I, Ramirez L, Altisent C, Parra R, Vidal F. Rapid molecular diagnosis of von Willebrand disease by direct sequencing. Detection of 12 novel putative mutations in VWF gene. Thrombosis and haemostasis. 2009;101(3):570-576.

143. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008;452(7189):872-876.

144. International Human Genome Sequencing C. Finishing the euchromatic sequence of the human genome. Nature. 2004;431(7011):931-945.

145. Levy SE, Myers RM. Advancements in Next-Generation Sequencing. Annual review of genomics and human genetics. 2016;17(95-115.

146. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nature reviews Genetics. 2016;17(6):333-351.

147. Metzker ML. Sequencing technologies - the next generation. Nature reviews Genetics. 2010;11(1):31-46.

148. Liu L, Li Y, Li S, et al. Comparison of next-generation sequencing systems. Journal of biomedicine & biotechnology. 2012;2012(251364.

149. Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008;456(7218):53-59.

150. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. Molecular cell. 2015;58(4):586-597.

151. Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. Science. 2009;323(5910):133-138.

152. Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. Zero-mode waveguides for single-molecule analysis at high concentrations. Science. 2003;299(5607):682-686.

153. Wang Y, Yang Q, Wang Z. The evolution of nanopore sequencing. Frontiers in genetics. 2014;5(449.

154. Swystun LL, James PD. Genetic diagnosis in hemophilia and von Willebrand disease. Blood reviews. 2017;31(1):47-56.

155. Tjernberg P, Vos HL, Castaman G, Bertina RM, Eikenboom JC. Dimerization and multimerization defects of von Willebrand factor due to mutated cysteine residues. Journal of thrombosis and haemostasis : JTH. 2004;2(2):257-265.

156. Bowen DJ, Collins PW. An amino acid polymorphism in von Willebrand factor correlates with increased susceptibility to proteolysis by ADAMTS13. Blood. 2004;103(3):941-947.

157. Casonato A, Pontara E, Sartorello F, et al. Reduced von Willebrand factor survival in type Vicenza von Willebrand disease. Blood. 2002;99(1):180-184.

158. Othman M, Chirinian Y, Brown C, et al. Functional characterization of a 13-bp deletion (c.-1522_-1510del13) in the promoter of the von Willebrand factor gene in type 1 von Willebrand disease. Blood. 2010;116(18):3645-3652.

159. Castaman G, Plate M, Giacomelli SH, Rodeghiero F, Duga S. Alterations of mRNA processing and stability as a pathogenic mechanism in von Willebrand factor quantitative deficiencies. Journal of thrombosis and haemostasis : JTH. 2010;8(12):2736-2742.

160. de Jong A, Eikenboom J. Von Willebrand disease mutation spectrum and associated mutation mechanisms. Thrombosis research. 2017;159(65-75.

161. Tjernberg P, Castaman G, Vos HL, Bertina RM, Eikenboom JC. Homozygous C2362F von Willebrand factor induces intracellular retention of mutant von Willebrand factor resulting in autosomal recessive severe von Willebrand disease. British journal of haematology. 2006;133(4):409-418.

162. Baronciani L, Federici AB, Cozzi G, et al. Expression studies of missense mutations p.D141Y, p.C275S located in the propeptide of von Willebrand factor in patients with type 3 von Willebrand disease. Haemophilia : the official journal of the World Federation of Hemophilia. 2008;14(3):549-555.

163. Schneppenheim R, Castaman G, Federici AB, et al. A common 253-kb deletion involving VWF and TMEM16B in German and Italian patients with severe von Willebrand disease type 3. Journal of thrombosis and haemostasis : JTH. 2007;5(4):722-728.

164. Gupta PK, Adamtziki E, Budde U, et al. Gene conversions are a common cause of von Willebrand disease. British journal of haematology. 2005;130(5):752-758.

165. Lyons SE, Bruck ME, Bowie EJ, Ginsburg D. Impaired intracellular transport produced by a subset of type IIA von Willebrand disease mutations. The Journal of biological chemistry. 1992;267(7):4424-4430.

166. Hassenpflug WA, Budde U, Obser T, et al. Impact of mutations in the von Willebrand factor A2 domain on ADAMTS13-dependent proteolysis. Blood. 2006;107(6):2339-2345.

167. Jacobi PM, Gill JC, Flood VH, Jakab DA, Friedman KD, Haberichter SL. Intersection of mechanisms of type 2A VWD through defects in VWF multimerization, secretion, ADAMTS-13 susceptibility, and regulated storage. Blood. 2012;119(19):4543-4553.

168. Rodeghiero F, Castaman G, Tosetto A. Optimizing treatment of von Willebrand disease by using phenotypic and molecular data. Hematology American Society of Hematology Education Program. 2009;113-123.

169. Grainick HR, Williams SB, McKeown LP, et al. Von Willebrand's disease with spontaneous platelet aggregation induced by an abnormal plasma von Willebrand factor. The Journal of clinical investigation. 1985;76(4):1522-1529.

170. Larsen DM, Haberichter SL, Gill JC, Shapiro AD, Flood VH. Variability in platelet- and collagen-binding defects in type 2M von Willebrand disease. Haemophilia : the official journal of the World Federation of Hemophilia. 2013;19(4):590-594.

171. Flood VH, Schlauderaff AC, Haberichter SL, et al. Crucial role for the VWF A1 domain in binding to type IV collagen. Blood. 2015;125(14):2297-2304.

172. Casonato A, Sartorello F, Cattini MG, et al. An Arg760Cys mutation in the consensus sequence of the von Willebrand factor propeptide cleavage site is responsible for a new von Willebrand disease variant. Blood. 2003;101(1):151-156.

173. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine : official journal of the American College of Medical Genetics. 2015;17(5):405-424.

174. Corrales I, Ramirez L, Ayats J, Altisent C, Parra R, Vidal F. Integration of molecular and clinical data of 40 unrelated von Willebrand Disease families in a Spanish locus-specific mutation database: first release including 58 mutations. Haematologica. 2010;95(11):1982-1984.

175. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols. 2009;4(7):1073-1081.

176. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic acids research. 2011;39(17):e118.

177. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nature methods. 2010;7(4):248-249.

178. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nature methods. 2010;7(8):575-576.

179. Pejaver V, Urresti J, Lugo-Martinez J, et al. MutPred2: inferring the molecular and phenotypic impact of amino acid variants. bioRxiv. 2017;

180. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. Journal of computational biology : a journal of computational molecular cell biology. 1997;4(3):311-323.

181. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. Journal of molecular biology. 1991;220(1):49-65.

182. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic acids research. 2009;37(9):e67.

183. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. Journal of computational biology : a journal of computational molecular cell biology. 2004;11(2-3):377-394.

184. McGrath RT, McRae E, Smith OP, O'Donnell JS. Platelet von Willebrand factor--structure, function and biological importance. British journal of haematology. 2010;148(6):834-843.

185. Cumming A, Grundy P, Keeney S, et al. An investigation of the von Willebrand factor genotype in UK patients diagnosed to have type 1 von Willebrand disease. Thrombosis and haemostasis. 2006;96(5):630-641.

186. Goodeve A, Eikenboom J, Castaman G, et al. Phenotype and genotype of a cohort of families historically diagnosed with type 1 von Willebrand disease in the European study, Molecular and Clinical Markers for the Diagnosis and Management of Type 1 von Willebrand Disease (MCMDM-1VWD). Blood. 2007;109(1):112-121.

187. Nesbitt IM, Hampton KK, Preston FE, Peake IR, Goodeve AC. A common splice site mutation is shared by two families with different type 2N von Willebrand disease mutations. Thrombosis and haemostasis. 1999;82(3):1061-1064.

188. Mertes G, Ludwig M, Finkelnburg B, et al. A G+3-to-T donor splice site mutation leads to skipping of exon 50 in von Willebrand factor mRNA. Genomics. 1994;24(1):190-191.

189. Cooper DN, Berg LP, Kakkar VV, Reiss J. Ectopic (illegitimate) transcription: new possibilities for the analysis and diagnosis of human genetic disease. Annals of medicine. 1994;26(1):9-14.

190. Pagliari MT, Baronciani L, Garcia Oya I, et al. A synonymous (c.3390C>T) or a splice-site (c.3380-2A>G) mutation causes exon 26 skipping in four patients with von Willebrand disease (2A/IIE). Journal of thrombosis and haemostasis : JTH. 2013;11(7):1251-1259.

191. Corrales I, Ramirez L, Altisent C, Parra R, Vidal F. The study of the effect of splicing mutations in von Willebrand factor using RNA isolated from patients' platelets and leukocytes. Journal of thrombosis and haemostasis : JTH. 2011;9(4):679-688.

192. Holbrook JA, Neu-Yilik G, Hentze MW, Kulozik AE. Nonsense-mediated decay approaches the clinic. Nature genetics. 2004;36(8):801-808.

193. Sadler JE. von Willebrand factor in its native environment. Blood. 2013;121(14):2583-2584.

194. Booyse FM, Quarfoot AJ, Chediak J, Stemerman MB, Maciag T. Characterization and properties of cultured human von Willebrand umbilical vein endothelial cells. Blood. 1981;58(4):788-796.

195. Levene RB, Booyse FM, Chediak J, Zimmerman TS, Livingston DM, Lynch DC. Expression of abnormal von Willebrand factor by endothelial cells from a patient with type IIA von Willebrand disease. Proceedings of the National Academy of Sciences of the United States of America. 1987;84(18):6550-6554.

196. Ewenstein BM, Inbal A, Pober JS, Handin RI. Molecular studies of von Willebrand disease: reduced von Willebrand factor biosynthesis, storage, and release in endothelial cells derived from patients with type I von Willebrand disease. Blood. 1990;75(7):1466-1472.

197. Ormiston ML, Toshner MR, Kiskin FN, et al. Generation and Culture of Blood Outgrowth Endothelial Cells from Human Peripheral Blood. Journal of visualized experiments : JoVE. 2015;106):e53384.

198. Bowman ML, Pluthero FG, Tuttle A, et al. Discrepant platelet and plasma von Willebrand factor in von Willebrand disease patients with p.Pro2808Leufs*24. Journal of thrombosis and haemostasis : JTH. 2017;15(7):1403-1411.

199. Lin Y, Weisdorf DJ, Solovey A, Hebbel RP. Origins of circulating endothelial cells and endothelial outgrowth from blood. The Journal of clinical investigation. 2000;105(1):71-77.

200. Starke RD, Paschalaki KE, Dyer CE, et al. Cellular and molecular basis of von Willebrand disease: studies on blood outgrowth endothelial cells. Blood. 2013;121(14):2773-2784.

201. Wang JW, Bouwens EA, Pintao MC, et al. Analysis of the storage and secretion of von Willebrand factor in blood outgrowth endothelial cells derived from patients with von Willebrand disease. Blood. 2013;121(14):2762-2772.

202. James PD, Notley C, Hegadorn C, et al. The mutational spectrum of type 1 von Willebrand disease: Results from a Canadian cohort study. Blood. 2007;109(1):145-154.

203. Johansson AM, Hallden C, Sall T, Lethagen S. Variation in the VWF gene in Swedish patients with type 1 von Willebrand Disease. Annals of human genetics. 2011;75(4):447-455.

204. Robertson JD, Yenson PR, Rand ML, et al. Expanded phenotype-genotype correlations in a pediatric population with type 1 von Willebrand disease. Journal of thrombosis and haemostasis : JTH. 2011;9(9):1752-1760.

205. Meyer D, Fressinaud E, Gaucher C, et al. Gene defects in 150 unrelated French cases with type 2 von Willebrand disease: from the patient to the gene. INSERM Network on Molecular Abnormalities in von Willebrand Disease. Thrombosis and haemostasis. 1997;78(1):451-456.

206. Federici AB, Mannucci PM, Castaman G, et al. Clinical and molecular predictors of thrombocytopenia and risk of bleeding in patients with von Willebrand disease type 2B: a cohort study of 67 patients. Blood. 2009;113(3):526-534.

207. Ahmad F, Jan R, Kannan M, et al. Characterisation of mutations and molecular studies of type 2 von Willebrand disease. Thrombosis and haemostasis. 2013;109(1):39-46.

208. JAMES PD, NOTLEY C, HEGADORN C, et al. Challenges in defining type 2M von Willebrand disease: results from a Canadian cohort study. Journal of Thrombosis and Haemostasis. 2007;5(9):1914-1922.

209. Zhang ZP, Blomback M, Egberg N, Falk G, Anvret M. Characterization of the von Willebrand factor gene (VWF) in von Willebrand disease type III patients from 24 families of Swedish and Finnish origin. Genomics. 1994;21(1):188-193.

210. Gupta PK, Saxena R, Adamtziki E, et al. Genetic defects in von Willebrand disease type 3 in Indian and Greek patients. Blood cells, molecules & diseases. 2008;41(2):219-222.

211. Sutherland MS, Keeney S, Bolton-Maggs PH, Hay CR, Will A, Cumming AM. The mutation spectrum associated with type 3 von Willebrand disease in a cohort of patients from the north west of England. Haemophilia : the official journal of the World Federation of Hemophilia. 2009;15(5):1048-1057.

212. Baronciani L, Cozzi G, Canciani MT, et al. Molecular defects in type 3 von Willebrand disease: updated results from 40 multiethnic patients. Blood cells, molecules & diseases. 2003;30(3):264-270.

213. Bowman M, Tuttle A, Notley C, et al. The genetics of Canadian type 3 von Willebrand disease: further evidence for co-dominant inheritance of mutant alleles. Journal of thrombosis and haemostasis : JTH. 2013;11(3):512-520.

214. Federici AB, Bucciarelli P, Castaman G, et al. Management of inherited von Willebrand disease in Italy: results from the retrospective study on 1234 patients. Seminars in thrombosis and hemostasis. 2011;37(5):511-521.

215. Yadegari H, Driesen J, Pavlova A, Biswas A, Hertfelder HJ, Oldenburg J. Mutation distribution in the von Willebrand factor gene related to the different von

Willebrand disease (VWD) types in a cohort of VWD patients. Thrombosis and haemostasis. 2012;108(4):662-671.

216. Veyradier A, Boisseau P, Fressinaud E, et al. A Laboratory Phenotype/Genotype Correlation of 1167 French Patients From 670 Families With von Willebrand Disease: A New Epidemiologic Picture. Medicine. 2016;95(11):e3038.
217. Lillicrap D. von Willebrand disease: advances in pathogenetic understanding, diagnosis, and therapy. Blood. 2013;122(23):3735-3740.
218. Flood VH, Gill JC, Morateck PA, et al. Common VWF exon 28 polymorphisms in African Americans affecting the VWF activity assay by ristocetin cofactor. Blood. 2010;116(2):280-286.
219. Montgomery RR, Flood VH. What have we learned from large population studies of von Willebrand disease? Hematology American Society of Hematology Education Program. 2016;2016(1):670-677.
220. Flood VH, Gill JC, Morateck PA, et al. Gain-of-function GPIb ELISA assay for VWF activity in the Zimmerman Program for the Molecular and Clinical Biology of VWD. Blood. 2011;117(6):e67-74.
221. Bodo I, Eikenboom J, Montgomery R, et al. Platelet-dependent von Willebrand factor activity. Nomenclature and methodology: communication from the SSC of the ISTH. Journal of thrombosis and haemostasis : JTH. 2015;13(7):1345-1350.
222. Corrales I, Catarino S, Ayats J, et al. High-throughput molecular diagnosis of von Willebrand disease by next generation sequencing methods. Haematologica. 2012;97(7):1003-1007.
223. Genomes Project C, Abecasis GR, Altshuler D, et al. A map of human genome variation from population-scale sequencing. Nature. 2010;467(7319):1061-1073.
224. Grossmann V, Kohlmann A, Eder C, et al. Molecular profiling of chronic myelomonocytic leukemia reveals diverse mutations in >80% of patients with TET2 and EZH2 being of high prognostic relevance. Leukemia. 2011;25(5):877-879.
225. Liang Q, Qin H, Ding Q, et al. Molecular and clinical profile of VWD in a large cohort of Chinese population: application of next generation sequencing and CNVplex((R)) technique. Thromb Haemost. 2017;117(8):1534-1548.
226. Federici AB, Canciani MT. Clinical and laboratory versus molecular markers for a correct classification of von Willebrand disease. Haematologica. 2009;94(5):610-615.
227. Federici AB, Bucciarelli P, Castaman G, et al. The bleeding score predicts clinical outcomes and replacement therapy in adults with von Willebrand disease. Blood. 2014;123(26):4037-4044.
228. Yohe S, Thyagarajan B. Review of Clinical Next-Generation Sequencing. Archives of pathology & laboratory medicine. 2017;141(11):1544-1557.

229. Favaloro EJ. Genetic testing for von Willebrand disease: the case against. Journal of thrombosis and haemostasis : JTH. 2010;8(1):6-12.

230. Flood VH, Gill JC, Christopherson PA, et al. Critical von Willebrand factor A1 domain residues influence type VI collagen binding. Journal of thrombosis and haemostasis : JTH. 2012;10(7):1417-1424.

231. Yadegari H, Driesen J, Hass M, Budde U, Pavlova A, Oldenburg J. Large deletions identified in patients with von Willebrand disease using multiple ligation-dependent probe amplification. Journal of thrombosis and haemostasis : JTH. 2011;9(5):1083-1086.

232. Bernardi F, Patracchini P, Gemmati D, et al. In-frame deletion of von Willebrand factor A domains in a dominant type of von Willebrand disease. Human molecular genetics. 1993;2(5):545-548.

233. Kenter AL, Birshtein BK. Chi, a promoter of generalized recombination in lambda phage, is present in immunoglobulin genes. Nature. 1981;293(5831):402-404.

234. Eikenboom JC, Vink T, Briet E, Sixma JJ, Reitsma PH. Multiple substitutions in the von Willebrand factor gene that mimic the pseudogene sequence. Proceedings of the National Academy of Sciences of the United States of America. 1994;91(6):2221-2224.

235. Castaman G, Lethagen S, Federici AB, et al. Response to desmopressin is influenced by the genotype and phenotype in type 1 von Willebrand disease (VWD): results from the European Study MCMDM-1VWD. Blood. 2008;111(7):3531-3539.

236. Kasatkar P, Shetty S, Ghosh K. Genetic heterogeneity in a large cohort of Indian type 3 von Willebrand disease patients. PloS one. 2014;9(3):e92575.

237. Jorieux S, Gaucher C, Goudemand J, Mazurier C. A novel mutation in the D3 domain of von Willebrand factor markedly decreases its ability to bind factor VIII and affects its multimerization. Blood. 1998;92(12):4663-4670.

238. Casana P, Martinez F, Haya S, Lorenzo JI, Espinos C, Aznar JA. Q1311X: a novel nonsense mutation of putative ancient origin in the von Willebrand factor gene. British journal of haematology. 2000;111(2):552-555.

239. Johnsen JM, Auer PL, Morrison AC, et al. Common and rare von Willebrand factor (VWF) coding variants, VWF levels, and factor VIII levels in African Americans: the NHLBI Exome Sequencing Project. Blood. 2013;122(4):590-597.

240. Bellissimo DB, Christopherson PA, Flood VH, et al. VWF mutations and new sequence variations identified in healthy controls are more frequent in the African-American population. Blood. 2012;119(9):2135-2140.

241. Flood VH. New insights into genotype and phenotype of VWD. Hematology American Society of Hematology Education Program. 2014;2014(1):531-535.

242. Castaman G, Eikenboom JC, Bertina RM, Rodeghiero F. Inconsistency of association between type 1 von Willebrand disease phenotype and genotype in

families identified in an epidemiological investigation. Thrombosis and haemostasis. 1999;82(3):1065-1070.

243. Casana P, Martinez F, Haya S, Espinos C, Aznar JA. Significant linkage and non-linkage of type 1 von Willebrand disease to the von Willebrand factor gene. British journal of haematology. 2001;115(3):692-700.

244. James PD, Paterson AD, Notley C, et al. Genetic linkage and association analysis in type 1 von Willebrand disease: results from the Canadian type 1 VWD study. Journal of thrombosis and haemostasis : JTH. 2006;4(4):783-792.

245. Eikenboom J, Van Marion V, Putter H, et al. Linkage analysis in families diagnosed with type 1 von Willebrand disease in the European study, molecular and clinical markers for the diagnosis and management of type 1 VWD. Journal of thrombosis and haemostasis : JTH. 2006;4(4):774-782.

246. Smith NL, Chen MH, Dehghan A, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. Circulation. 2010;121(12):1382-1392.

247. van Loon J, Dehghan A, Weihong T, et al. Genome-wide association studies identify genetic loci for low von Willebrand factor levels. European journal of human genetics : EJHG. 2016;24(7):1035-1040.

248. Stenson PD, Mort M, Ball EV, et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Human genetics. 2017;136(6):665-677.

249. Hampshire DJ, Goodeve AC. The international society on thrombosis and haematosis von Willebrand disease database: an update. Seminars in thrombosis and hemostasis. 2011;37(5):470-479.

250. Johnston JJ, Biesecker LG. Databases of genomic variation and phenotypes: existing resources and future needs. Human molecular genetics. 2013;22(R1):R27-31.

251. Peterson TA, Doughty E, Kann MG. Towards precision medicine: advances in computational approaches for the analysis of human variants. Journal of molecular biology. 2013;425(21):4047-4063.

252. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nature reviews Genetics. 2002;3(4):285-298.

253. Krawczak M, Reiss J, Cooper DN. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Human genetics. 1992;90(1-2):41-54.

254. James PD, O'Brien LA, Hegadorn CA, et al. A novel type 2A von Willebrand factor mutation located at the last nucleotide of exon 26 (3538G>A) causes skipping of 2 nonadjacent exons. Blood. 2004;104(9):2739-2745.

255. Martorell L, Luce E, Vazquez JL, et al. Advanced cell-based modeling of the royal disease: characterization of the mutated F9 mRNA. Journal of thrombosis and haemostasis : JTH. 2017;15(11):2188-2197.

256. Yadegari H, Biswas A, Akhter MS, et al. Intron retention resulting from a silent mutation in the VWF gene that structurally influences the 5' splice site. Blood. 2016;128(17):2144-2152.

257. Hawke L, Bowman ML, Poon MC, Scully MF, Rivard GE, James PD. Characterization of aberrant splicing of von Willebrand factor in von Willebrand disease: an underrecognized mechanism. Blood. 2016;128(4):584-593.

258. Rehm HL, Bale SJ, Bayrak-Toydemir P, et al. ACMG clinical laboratory standards for next-generation sequencing. Genetics in medicine : official journal of the American College of Medical Genetics. 2013;15(9):733-747.

259. Borràs N, Ramírez L, Comes N, et al. Development of a next generation sequencing panel for diagnostic and investigation of bleeding disorders. European Society of Human Genetics; 2016; Barcelona, Spain; 2016.

260. James PD, Lillicrap D. von Willebrand disease: clinical and laboratory lessons learned from the large von Willebrand disease studies. American journal of hematology. 2012;87 Suppl 1(S4-11.

261. Sanders YV, van der Bom JG, Isaacs A, et al. CLEC4M and STXBP5 gene variations contribute to von Willebrand factor level variation in von Willebrand disease. Journal of thrombosis and haemostasis : JTH. 2015;13(6):956-966.

262. Holm IA, Agrawal PB, Ceyhan-Birsoy O, et al. The BabySeq project: implementing genomic sequencing in newborns. BMC pediatrics. 2018;18(1):225.

263. Lindor NM, Thibodeau SN, Burke W. Whole-Genome Sequencing in Healthy People. Mayo Clinic proceedings. 2017;92(1):159-172.

264. Swystun LL, James P. Using genetic diagnostics in hemophilia and von Willebrand disease. Hematology American Society of Hematology Education Program. 2015;2015(152-159.