



This is a repository copy of *Molecular diagnosis of von Willebrand disease.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/113478/>

Version: Accepted Version

---

**Article:**

Baronciani, L., Goodeve, A. and Peyvandi, F. (2017) Molecular diagnosis of von Willebrand disease. *Haemophilia*, 23 (2). pp. 188-197. ISSN 1351-8216

<https://doi.org/10.1111/hae.13175>

---

This is the peer reviewed version of the following article: Baronciani, L., Goodeve, A. and Peyvandi, F. (2017), Molecular diagnosis of von Willebrand disease. *Haemophilia*. which has been published in final form at <https://doi.org/10.1111/hae.13175>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

## **Molecular Diagnosis of Von Willebrand Disease**

Luciano Baronciani\*, Anne Goodeve‡ and Flora Peyvandi\*†

\*Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico and Luigi Villa Foundation; †Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy  
‡Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield Medical School, Sheffield, UK.

**Address for Correspondence:** Flora Peyvandi

Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico and Luigi Villa Foundation; Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy  
Via Pace 9, 20122 Milan Italy  
Tel: +39 02 5503 4456; Fax: +39 02 5410 0125; e-mail: flora.peyvandi@unimi.it

**Running Title:** Molecular Characterization of VWD

**Key Words:** von Willebrand disease, von Willebrand factor, molecular characterization, DNA sequence analysis, Next Generation Sequence analysis, NGS, Multiplex Ligation-dependent Probe Amplification.

**Abstract word count:** 212

**Text word count:** 3781

**Number of Tables and Figures:** Tables 3, Figures 3

## **Abstract**

The role of molecular characterization in the diagnosis of VWD is not essential if the patients have been extensively investigated using phenotypic analysis. On the other hand, if some of these phenotype assays are not available, the identification of the mutation causing the disease could be crucial for an accurate diagnosis. Nevertheless, there are several reasons for performing molecular analysis in patients phenotypically well characterized, e.g. to identify the mutation causing VWD can be useful for patients and their family members when prenatal diagnosis is required (type 3 or severe type 2).

In this manuscript we report the techniques used for the molecular characterization of suspected VWD patients. We describe the use of online VWF database and online SNV (single nucleotide variation) databases, the former to verify whether a candidate mutation has been previously identified in other VWD patients and the latter to ascertain whether a putative mutation has been reported earlier in healthy individuals. We listed the available *in silico* analysis tools, to determine the predicted pathogenicity of a sequence variant and to establish its possible negative effect on the normal splicing process. We also report the strategy that can be used to identify VWD type 2 patients' mutations in subjects that have been fully characterized using the phenotype assays.

## Introduction

### *The role of molecular characterization in VWD diagnosis*

The importance of molecular characterization in the diagnosis of VWD patients is inversely proportional to the number and quality of phenotypic assays run in the laboratory. If all VWF assays are used [1], the identification of the molecular defects will only confirm the patient's phenotypic diagnosis. On the other hand, if some of these phenotypic assays are not available, e.g. the VWF-FVIII binding assay (VWF:FVIII B), ristocetin-induced platelet agglutination (RIPA) or multimer analysis, then the identification of the mutation causing VWD will be crucial for the patient's diagnosis.

The VWF:FVIII B assay can provide differential diagnosis between mild hemophilia and VWD type 2N (Normandy). However, the assay is not available in many laboratories, therefore DNA sequence analysis for the VWF gene (*VWF*) exons encoding the D'D3 domains can be a valuable alternative to the use of VWF:FVIII B assay.

VWD types 2B and 2A (IIA) patients can present similar phenotypic parameters. The RIPA assay, performed using platelet-rich plasma, evaluates the affinity of VWF to bind to the glycoprotein (GP)Ib $\alpha$  platelet receptor, allowing differential diagnosis between VWD types 2B and 2A (IIA). However, if the RIPA assay is not available or only patient frozen plasma is available, DNA sequence analysis of the VWF exon encoding the A1 domain (type 2B) or A2 domain (type 2A/IIA) can be used to establish the patient diagnosis. Even in the presence of an enhanced RIPA, outcome of the DNA sequence analysis should be performed because the 2B-like phenotype can also be caused (although less frequently) by gain-of-function mutations in the GPIb $\alpha$  gene (*GP1BA*). The identification of mutations in the VWF A1 domain (VWD type 2B) or in the *GP1BA* (Pseudo-VWD) is a valid alternative to platelet-mixing tests [1] required for the differential diagnosis.

Also in the case of VWD types 2A (IIA) and 2M patients can present similar phenotypic parameters (i.e. markedly reduced VWF:RC<sub>0</sub> or other platelet-dependent VWF activity

determination in comparison to VWF:Ag level and decreased RIPA) and the phenotypic diagnosis might be difficult if multimer analysis is not available. Once again the molecular analysis can be used to obtain, in the large majority of the cases, the patients diagnosis by investigating exon 28 encoding the A1 domain (VWD type 2M) or the A2 domain (VWD type 2A/IIA).

Nevertheless, there are several reasons for also performing molecular analysis in well phenotypically characterized diagnosed patients. VWD type 3 diagnosis can be performed using a reliable VWF:Ag assay. However, the identification of carrier status among family members might be uncertain using the phenotypic approach, whereas it can be easily ascertained once the proband's mutation(s) have been identified.

The identification of homozygous large deletions in VWD type 3 is an important risk factor for the development of anti-VWF alloantibodies [2,3] and these defects should always be investigated in previously untreated patients. Prenatal diagnosis (PND) or pre-implantation diagnosis in VWD type 3, particularly in cases where the parents already have a severely affected child, may be required and can be based on analysis of DNA from chorionic villi/ amniocytes or embryo cells. In addition, in families with severe VWD types 1 or 2, the molecular approach can be easily used to perform post-natal analysis [4-6].

*The von Willebrand factor, gene, pseudogene, mRNA transcription, pro-VWF functional domains and its complex biosynthesis.*

The *VWF* has been localized to chromosome 12, containing 52 exons with a length of about 178 kilobases (kb) [7]. A partial unprocessed VWF pseudogene (*VWFP1*, exons 23-34) with 97% homology with the gene has been mapped to chromosome 22 [8]. Transcription of *VWF* results in a mRNA of approximately 8.7 kb in size. VWF is synthesized by endothelial cells and megakaryocytes as a precursor polypeptide of 2813 amino acids (pre-pro-VWF), including a 22-residue signal peptide, 741-residue propeptide

and 2050-residue mature subunit [7]. The pro-VWF (260 kD) is organized into repeats of homologous domains (D1-D2-D'-D3-A1-A2-A3-D4-C1-C2-C3-C4-C5-C6-CK) [9,10] for which specific functions have been identified (Figure 1). The propeptide, D1-D2 domains have a role in the multimerization process and VWF localization to storage organelles, the D'-D3 domains exhibit a binding site for factor VIII (FVIII) and plays an important role in VWF multimerization. The A1 domain is the only known binding site for the GPIIb $\alpha$  platelet receptor, in addition, the A1 domain also provides a binding site for both type IV and VI collagen. The A2 domain contains the cleavage site for ADAMTS13 (a disintegrin-like and metalloprotease with thrombospondin type 1 motif, member 13) [11]. The A3 domain is the binding site for fibrillar collagen types I and III. The C4 domain which includes the RGD sequence, is the binding site for the integrin  $\alpha$ IIb $\beta$ 3 (GPIIb/IIIa). A complex post-translational biosynthesis is required in order to obtain multimeric VWF. After translocation into the endoplasmic reticulum, the signal peptide is cleaved and the pro-VWF dimerizes through disulfide bonds between C-terminal cysteine residues of the CK domain [12]. Multimerization continues in the Golgi apparatus where dimers, through disulfide bonds at the N-terminal region of the subunits (D' and D3 domain), yield multimers that may exceed 20,000 kD. The propeptide, that is cleaved prior to protein secretion has a key role in the assembly of VWF multimers [13] and to traffic the VWF into the storage organelles (Weibel-Palade bodies of endothelial cells or platelet  $\alpha$ -granules) [14,15]. An update on the complex process of VWF biosynthesis, secretion and clearance has been reported recently by Lenting et al. [16].

## **Techniques used in the molecular characterization of von Willebrand disease patients**

*The use of genomic DNA versus mRNA*

Usually, a mutation search is performed using genomic DNA, comprising either PCR amplification of *VWF* exons followed by Sanger sequencing or more recently, next generation DNA sequencing (NGS). Messenger RNA (mRNA) isolated from patient's platelets can also be used. Nevertheless, the instability of the latter samples and the more laborious techniques required to obtain it often reduces the use of mRNA analysis. This strategy is usually applied to assess possible splicing mutations or intronic mutations identified in possible regulatory areas. The study of mRNA can be undertaken in VWD patients where no mutations have been previously identified after evaluation performed at the exonic DNA level. This approach is reasonable in VWD types 3 or 2 patients (with a proven family bleeding history), but not in VWD type 1 where mutations are not always identified. In this case, an alternative mRNA sequence due to altered splicing resulting from a possible deep intronic sequence change, may be found at the mRNA level. The identification of altered splicing is not always possible since mRNA decay in VWD patients presenting premature termination codons has been previously reported [17-19].

*Direct DNA sequence (Sanger) analysis.*

The Sanger method allows establishing the nucleotide sequence of a DNA fragment (direct DNA sequence analysis). This technique is the most frequently adopted for the identification of *VWF* mutations. However, the new technology of NGS is reducing the use of the Sanger sequencing. Currently, the cost of NGS still makes the use of Sanger sequencing convenient in those cases where the mutation search is localized to specific *VWF* exons (i.e. most common VWD type 2 variants). In addition, this new technique is not yet widely available and therefore its use is limited to the more advanced genetic laboratories. For the investigation of the full *VWF* coding region along with the intronic flanking regions, the 5' and 3' untranslated regions and a portion of the promoter, using Sanger sequencing is extremely time-consuming, mainly due to the large size of the

coding region of the gene itself (8442 bp). This has impaired the molecular characterization of patients with VWD types 1 and 3 in comparison with type 2 where mutation analysis is often performed for specific VWF regions. Direct sequencing is available in many laboratories and is particularly indicated to identify the gene mutation in VWD type 2 variants (see VWD type 2 variant analysis). However, Sanger sequencing is not able to identify large heterozygous deletions or duplications.

#### *Next Generation Sequence (NGS) analysis*

Newly developed genetic methods such as NGS are becoming faster and progressively cheaper, allowing wider study of *VWF* mutations at a very reasonable cost [20]. NGS enables investigation of the entire *VWF* coding region in a short time and therefore this approach is better suited to investigate VWD types 1 and 3 variants. When possible, the study of type 2 variants can also be undertaken using this technique. However, patients DNA samples should be sent to genetic laboratories that already have experience with this technique in order to contain the cost of the assay. Analysis can be performed using either PCR amplification or sequence capture through hybridisation to prepare the template for NGS. All primers must be specifically designed to avoid amplification of *VWFP1* and also to prevent the presence of common single nucleotide variants (SNV) within PCR primer sequences. This technique allows investigation of *VWF* in a relatively short time of many patients. It is currently only available in larger genetic laboratories, but its adoption is spreading. Use of a dosage analysis algorithms can enable detection of heterozygous large deletions and duplications and heterozygous gene conversions can also be detected. Candidate causative variants found with this technique require the use of Sanger sequencing to confirm the identified mutations. Recently, a large VWD cohort from Spain has been investigated using this approach [21]. The NGS analysis comprises three main phases: preparation of the DNA library, clonal amplification and cyclic array sequencing.



Each phase can be done using more than one method. These authors used a modified version of the NGS method previously reported by Corrales et al. [20] and the reliability of this NGS approach has already been confirmed by Fidalgo and colleagues [22] on a cohort of Portuguese VWD patients. In Figure 2 a schematic representation of this NGS procedure is summarized.

#### *Multiplex Ligation-dependent Probe Amplification (MLPA)*

Another technique, important in the molecular characterization of VWD variants, is MLPA analysis [23]. Large gene deletions and more rarely, large duplications may be present in *VWF*. Large gene deletions have been reported in all 3 disease types of VWD. If a deletion affects only one allele, PCR based techniques are not able to identify the mutation. The semi-quantitative approach of the MLPA technique (MRC-Holland, The Netherlands) can identify large gene deletions or duplications even when these defects are heterozygous [24]. Most large deletions that have been identified have been seen in VWD type 3 patients, however, in-frame large deletions have also been identified in VWD types 1 and 2. This technique which can be adopted in all laboratories that are able to perform Sanger sequencing enables the identification of large gene deletions or duplications of an exon or more.

#### **Online VWF databases**

##### *Pathological mutations identified in VWF*

An online database on *VWF* is available through the Scientific and Standardization Committee on VWF of the International Society of Thrombosis and Haemostasis at the site of the University of Sheffield (<http://www.vwf.group.shef.ac.uk/> ; accessed November 2016). The *VWF* mutation listing for the database is hosted through the European Association for Haemophilia and Allied Disorders (EAHAD) Coagulation Factor Variant

Databases portal ([https://grenada.lumc.nl/LOVD2/VWF/home.php?select\\_db=VWF](https://grenada.lumc.nl/LOVD2/VWF/home.php?select_db=VWF) , accessed November 2016). When missense, synonymous or splice site mutations are found, it is particularly important to verify whether they have previously been identified in other VWD patients with a similar phenotype, supporting the pathological role of the mutation. It is possible to contribute to these VWD databases by sending the new mutations identified along with information regarding patients' phenotype to the following address: <https://grenada.lumc.nl/LOVD2/VWF/submit.php> or <http://www.eahad-db.org/index.php>

### *Sequence variants identified in VWF*

There are now a number of databases that list SNV and their frequency within a number of different population groups. These include the single nucleotide polymorphism database, dbSNP ([www.ncbi.nlm.nih.gov/SNP](http://www.ncbi.nlm.nih.gov/SNP)), the 1000 Genomes database ([www.1000genomes.org](http://www.1000genomes.org)), Exome Aggregation Consortium, ExAC (<http://exac.broadinstitute.org>) and Exome Variant Server, EVS (<http://evs.gs.washington.edu/EVS>), accessed November 2016. ExAC is particularly useful as it lists sequence variants in approximately 60,000 individuals from several different populations. The position of *VWF* SNV should be taken into consideration during the design of oligonucleotides that will be used in any of the techniques mentioned above, in order to avoid mis-amplification. The use of these databases is useful in order to verify if the same missense, synonymous or candidate splice site mutation has been already identified in several healthy individuals, helping to exclude a pathological role of the identified variant.

### ***In silico analysis***

Newly developed genetic methods are becoming utilized in many laboratories, allowing wider molecular studies in many different diseases. Most of these software tools can be freely accessed online and this has resulted in a rapid increase in the detection of rare or novel sequence variations. The use of *in silico* analysis represents a valuable tool to determine the predicted pathogenicity of exonic and intronic sequence variants within a limited timescale and budget [25] [26].

#### *Splice site prediction software*

Several different software tools are available to assess the likely impact of a mutation on mRNA splicing, examples are listed in Table 1. Use of at least 3 of the recommended predictions tools is reported to reach a consensus result. Where the tools do not agree, further predictions should be undertaken to try and reach a consensus.

#### *Amino acid prediction software*

The potential damaging effect on protein structure or function of a single nucleotide variant can be evaluated using the *in silico* tools listed in Table 2. Tools should be selected that rely on different types of algorithm to analyse potential pathogenicity. A minimum of about 5 tools should be used to seek a consensus on likelihood of pathogenicity or neutrality.

### **Molecular characterization of VWD patients**

As previously mentioned, the role of molecular characterization in the diagnosis of VWD is not essential if the patients have been extensively investigated using phenotype analysis. However, identifying the mutations causing VWD can be useful for patients and their families in cases where prenatal diagnosis is required. In the following paragraphs, we

report the strategy that can be used to identify patients' mutations in subjects that have been fully characterized using the phenotype assays.

#### *Molecular characterization of type 1 & 3 VWD patients*

Once the patients' diagnosis as VWD types 1 or 3 has been undertaken using phenotypic methods, molecular analysis can be performed. The approach of investigating gene mutations causing types 1 and 3 is similar since in contrast to VWD type 2, where the large majority of mutations are located in exon 28 (see Figure 1), there is no particular area where mutations are identified. The whole *VWF* coding region should be investigated along with the intronic flanking regions, the 5' and 3' untranslated regions and the proximal *VWF* promoter. For VWD type 3, that segregates as a recessive disease, both alleles may harbour different mutations (compound heterozygous) or 2 copies of the same mutation (homozygous). For VWD type 1, that mainly segregates as a dominant disease, only one allele is affected, although there are reports regarding VWD type 1 due to compound heterozygous mutations [22,27]. In addition, the identification of both mutations in VWD type 3 patients is possible in almost all cases, whereas in VWD type 1 [27,28], mutations are not identified in almost 30% of the investigated patients (mainly in those where the VWF:Ag level is above 30 IU/dL). Moreover, the mutations identified in VWD type 3 are often clearly disease-causing mutations (nonsense, frame-shift and large deletions), whereas in VWD type 1, missense mutations often identified can only be considered candidate mutations due to the lack of firm evidence of their causal effect. **Figure 3** shows an algorithm for molecular characterization in patients affected with VWD types 1 and 3.

#### *Prenatal diagnosis (PND) of VWD type 3*

Compared with hemophilia, most patients with VWD show relatively mild bleeding symptoms. Therefore, PND is not usually required, except for occasional type 2 cases and

VWD type 3 families with a severe clinical history. PND is mainly required in cases where both parents are already known to be carriers of VWD type 3, with gene defects identified in their first affected child with severe bleeding symptoms. In the past, PND was performed using PCR of short tandem repeats (STR) in *VWF* [29]. To detect the gene mutations, the above mentioned sequencing and dosage analysis techniques can be used. For known mutations, sequence analysis is the first method of choice, whereas if the gene mutations are not known, NGS and if necessary, MLPA can be used [21] [24].

#### *Molecular characterization of VWD type 2 patients*

Differently to VWD types 1 and 3, the molecular diagnosis of type 2 variants can be performed in many laboratories. Characterization of VWD type 2 mutations does not require the study of the whole *VWF* gene, but only the evaluation of the exons encoding the specific functional domain(s) identified using the phenotypic approach. This strategy can be applied to all type 2 VWD (2A, 2B, 2M and 2N), although the evaluation of type 2A can be more complex due to the involvement of different functional domains implicated in the phenotype of these patients (see Figure 1).

In VWD type 2B that segregates as dominant disease, only one mutation is expected. These mutations lie in the A1 domain, between p.Glu1260 and p.Gly1479 and are associated with spontaneous binding of VWF to the platelet GPIIb/IIIa receptor (RIPA <0.7 mg/ml). To identify this defect, only the investigation of the 5' portion of exon 28 is necessary. Therefore one simple PCR, followed by sequence analysis can be used to identify the causative mutation. If no mutation is identified in the VWF A1 domain, in spite of a clearly enhanced RIPA result obtained from patient's platelet rich plasma, direct sequencing analysis of the *GP1BA* is necessary.

Pseudo-VWD, also known as platelet-type VWD, is due to missense mutations or an in-frame deletion in exon 2 of the *GP1BA*. Missense variants (between p.Trp246 and

p.Asp265) and an in-frame deletion; a 27-bp in-frame deletion Pro449\_Ser457del (1345\_1371del27) [30], cause GPIIb $\alpha$  to spontaneously bind to VWF resulting, as for VWD type 2B, in enhanced RIPA [31].

VWD type 2M also segregates as dominant disease. The missense mutations are more commonly located in the A1 domain. These defects are associated with reduced capacity of VWF to bind GPIIb $\alpha$  receptor (VWF:RCo/VWF:Ag <0.6). Therefore, to identify these defects, only the investigation of the 5' portion of exon 28 is necessary, similarly to VWD type 2B variants. However, collagen binding defects are also seen in VWD type 2M (VWF:CB/VWF:Ag <0.6) [32,33]. In this case, the majority of mutations are located in the A3 domain, and exons 29-32 should be investigated. A small proportion of A1 domain mutations result from collagen types IV or VI binding defects [34].

VWD type 2N segregates as a recessive disease, therefore both alleles are expected to be mutated (either as compound heterozygous or homozygous mutations). Homozygous mutations may be found in patients from the consanguineous partnerships, but also in European populations due to the high prevalence of the p.Arg854Gln mutation, often found in the homozygous form. Compound heterozygotes usually have a combination of a type 2N mutation with a type 1 or 3 VWF defect. Mutations affecting binding of VWF to FVIII (VWF:FVIII B) have been identified mainly in the first 272 amino acid residues of the mature subunit (domain D' and a portion of D3, encoded by exons 17-25), although mutations up to exon 26 have been reported to have mildly reduced VWF:FVIII B. Therefore, in this VWD type, identification of the mutations responsible for decreased VWF:FVIII B can be done by evaluating a relatively short region of VWF. However, for patients that are heterozygous for a type 2N missense mutation, identification of the second defect, requires the same approach used for VWD types 1 and 3 patients.

VWD type 2A patients have a heterogeneous group of mutations located in different VWF functional domains. All these mutations affect to different extents the VWF multimer

size, therefore these patients present a variable loss of high molecular weight (HMW) multimers.

The most common form of VWD type 2A segregates as dominant disease and the variants are located in the A2 domain, which contains the ADAMTS13 cleavage site. These variants, identified in the previous VWD classification as IIA [35], are characterized by the loss of high and intermediate sized multimers, associated with enhancement of the triplet bands due to increased susceptibility to ADAMTS13 proteolysis. Therefore, to identify these defects, only the investigation of the 3' portion of exon 28 is necessary.

The second most common form of type 2A VWD segregates as dominant disease and the variants are located in the D3 domain, which plays an important role in VWF multimerization. These variants, identified in the previous VWD classification as IIE [36], are characterized by a modest loss of HMW multimers, associated with the absence of satellite triplet bands due to reduced susceptibility to ADAMTS13 proteolysis. Therefore, to identify these defects, only the investigation of exons 22, 25-27 and the 5' end of exon 28 is necessary [37,38].

A rare form of VWD type 2A that also segregates as dominant disease is due to variants located in the CK domain, which plays an important role in VWF dimerization. These variants, identified in the previous VWD classification as IID [39], are characterized by the loss of HMW multimers, associated with the presence of "odd" band sizes in the intermediate resolution multimer analysis. Therefore, to identify these defects, investigation of only exons 51-52 is necessary [39].

A further rare form of VWD type 2A that segregates as a recessive disease, is due to variants located in the D1-D2 domains (propeptide) which play an important role in VWF multimerisation. The patients can be homozygous for a mutation in the D1 or D2 domains or compound heterozygous with a second defect elsewhere in *VWF*. These variants, identified in the previous VWD classification as IIC [36], are characterized by the loss of

high and intermediate molecular size multimers associated with the absence the satellite triplet bands due to reduced susceptibility to ADAMTS13 proteolysis. Therefore, to identify these defects, the investigation of exons 2-17 should be performed, although to date, mutations have been identified in the D1 domain, exons 6-7 and in exons 11-15 [40].

The above strategy to molecularly characterize the type 2 VWF variants is summarized in Table 3.

### **Conclusion**

VWD diagnosis can be achieved through both genetic analysis and phenotypic characterization of VWF. Sanger sequence analysis or NGS can identify mutations in most types 3 and 2 patients, but other etiology than *VWF* defects may contribute to reduced VWF levels in one third of type 1 VWD patients, since 30% of patients remain without mutations in the *VWF* protein coding sequence. Increasing use of NGS combined with dosage analysis derived from the NGS data will facilitate more complete analysis of *VWF* for mutations.



## Legends of the figures

Figure 1. Structure of the von Willebrand factor (VWF) precursor, its gene and pseudogene.

The schematic structure of the VWF precursor (prepro-VWF) consists of a signal peptide (residues 1–22), propeptide (residues 23–763), and mature subunit (residues 764–2813). The pro-VWF is organized into repeats of homologous structural domains (A, C and D). VWF binding sites for factor VIII, platelet glycoprotein (GP) Iba, collagen, platelet GPIIb/IIIa ( $\alpha$ IIb $\beta$ 3) and ADAMTS13 cleavage sites are shown. Arrows show the positions of mutations that cause von Willebrand disease (VWD) type 2. The dark arrows show the location of mutations that correspond to type 2A variants of VWD with characteristic multimer patterns because of increased proteolysis (IIA) or defective multimer assembly caused by mutations in the propeptide (IIC), the cysteine knot domain (IID), or the D3 domain (IIE). Recessive VWD type 2 variants are indicated by an \*.

Figure 2. A schematic representation of NGS analysis of the *VWF* as performed for the VWD patients by Batlle et al. [21].

Phase I) Library preparation was achieved using 61 pairs of primers to cover 1300 bp of the promoter region, the 52 *VWF* exons (including exon/intron boundaries) and 5' and 3' untranslated regions. The newly synthesized double stranded DNA obtained from the first polymerase chain reaction (PCR) was used as template in a second PCR. A new set of primers containing adapters for the NGS platform and a different index sequence for each sample were used in the second PCR in order to obtain the DNA library. Phase II) Library DNA is load into a flow cell in a MiSeq Desktop Sequencer (Illumina, San Diego, CA, USA) and clonal amplification is obtained by bridge PCR. DNA single strand fragments hybridized to linkers, through the sequence adapters, first with one end (a) and then with both ends and a bridge is formed (b). Isothermal extension is then initiated from the double

stranded ends of the bridge generating a complementary strand that is covalently attach to the linker (c). Double strand DNA denaturation causes displacement of original library DNA single strand (d). Repetition of steps (b) to (e) forms a cluster of clonally expanded sequences of single strand copies with either forward and reverse sequences (f). All reverse (or forward) single strand DNA fragments are removed by cleavage so that in each cluster all the molecules are identical (g) and ready to be sequenced. Phase III) Sequencing by synthesis. Sequence primers (blue) are added to the flow cell and hybridized to single stranded DNA (a). Primers are washed away and labeled/chemically modified nucleotides are added. Only a single nucleotide is incorporated in each cluster since the modified nucleotide cannot be extended (b). After washing of the flow cell and excitation of the incorporated labelled nucleotides, the emitted light is captured as an image identifying which of the four nucleotides have been incorporated in each cluster (c). In order to repeat this process the incorporated nucleotide is chemically unblocked and the labelled group removed (d). Repetition of steps (b) to (d) allows the progression of the sequencing reaction within each cluster.

Figure 3. The algorithm of molecular characterization of von Willebrand disease types 1 & 3 patients.

VWD, von Willebrand disease; MLPA, Multiplex Ligation-dependent Probe Amplification; SNV, single nucleotide variant; rVWF, recombinant von Willebrand factor

### ***Disclosures***

LB stated that he had no interests which might be perceived as posing a conflict or bias. AG has received honoraria for participating as a speaker at satellite symposia and educational meetings organized by Octapharma and Baxalta. She has received a sponsorship for the VWD mutation database from CSL Behring.

FP has received honoraria for participating as a speaker at educational meetings organized by Bayer, Grifols, Novo Nordisk, and Sobi. She is recipient of research grant funding from

Alexion, Biotest, Kedrion Biopharma, and Novo Nordisk paid to Fondazione Luigi Villa, and she has received consulting fees from Kedrion Biopharma, LFB and Octapharma. She is member of the Ablynx scientific advisory board.

### **Authorship**

LB wrote the manuscript. FP and AG critically revised the manuscript. All the authors approved the submitted and final version of the manuscript.

### References

1. Nichols WL, Hultin MB, James AH, Manco-Johnson MJ, Montgomery RR, Ortel TL, Rick ME, Sadler JE, Weinstein M, Yawn BP. von Willebrand disease (VWD): evidence-based diagnosis and management guidelines, the National Heart, Lung, and Blood Institute (NHLBI) Expert Panel report (USA). *Haemophilia* 2008; **14**: 171-232.
2. Shelton-Inloes BB, Chehab FF, Mannucci PM, Federici AB, Sadler JE. Gene deletions correlate with the development of alloantibodies in von Willebrand disease. *J Clin Invest* 1987; **79**: 1459-65.
3. Mannucci PM, Federici AB. Antibodies to von Willebrand factor in von Willebrand disease. In: Inhibitors to coagulation factors. Aledort LM, Hoyer LW, Reisner HM, White II GC, (eds). New York: Plenum Press, 1995; pp. 87-92.
4. Favaloro EJ. Genetic testing for von Willebrand disease: the case against. *J Thromb Haemost* 2010; **8**: 6-12.
5. Peake IR, Goodeve AC. Genetic testing for von Willebrand disease: the case for. *J Thromb Haemost* 2010; **8**: 13-6.
6. Mannucci PM. Genetic testing in von Willebrand disease: a rebuttal. *J Thromb Haemost* 2010; **8**: 860-7836.
7. Mancuso DJ, Tuley EA, Westfield LA, Worrall NK, Shelton-Inloes BB, Sorace JM, Alevy YG, Sadler JE. Structure of the gene for human von Willebrand factor. *J Biol Chem* 1989; **264**: 19514-27.
8. Mancuso DJ, Tuley EA, Westfield LA, Lester-Mancuso TL, Le Beau MM, Sorace JM, Sadler JE. Human von Willebrand factor gene and pseudogene: structural analysis and differentiation by polymerase chain reaction. *Biochemistry* 1991; **30**: 253-69.
9. 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**: 3164-71.
10. Zhou YF, Eng ET, Zhu J, Lu C, Walz T, Springer TA. Sequence and structure relationships within von Willebrand factor. *Blood* 2012; **120**: 449-58.

11. Chung DW, Fujikawa K. Processing of von Willebrand factor by ADAMTS-13. *Biochemistry* 2002; **41**: 11065-70.
12. Marti T, Roesselet S, Titani K, Walsh KA. Identification of disulfide-bridged substructures within human von Willebrand factor. *Biochemistry* 1987; **26**: 8099-109.
13. Verweij CL, Hart M, Pannekoek H. Expression of von Willebrand factor (vWF) cDNA in heterologous cells: requirement of the pro-polypeptide in vWF multimer formation. *EMBO J* 1987; **6**: 2885-90.
14. Haberichter SL, Fahs SA, Montgomery RR. Von Willebrand factor storage and multimerization: 2 independent intracellular processes. *Blood* 2003; **96**: 1808-15.
15. Haberichter SL, Jozwiak MA, Rosenberg JB, Christopherson PA, Montgomery RR. The von Willebrand factor propeptide (VWFpp) traffics an unrelated protein to storage. *Arterioscler Thromb Vasc Biol* 2002; **22**: 921-6.
16. Lenting PJ, Christophe OD, Denis CV. von Willebrand factor biosynthesis, secretion, and clearance: connecting the far ends. *Blood* 2015; **125**: 2019-28.
17. Plate M, Duga S, Baronciani L, La Marca S, Rubini V, Mannucci PM, Federici AB, Asselta R. Premature termination codon mutations in the von Willebrand factor gene are associated with allele-specific and position-dependent mRNA decay. *Haematologica* 2010; **95**: 172-4.
18. 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**: 2736-42.
19. 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**: 679-88.
20. Corrales I, Catarino S, Ayats J, Arteta D, Altisent C, Parra R, Vidal F. High-throughput molecular diagnosis of von Willebrand disease by next generation sequencing methods. *Haematologica* 2012; **97**: 1003-7.
21. Batlle J, Perez-Rodriguez A, Corrales I, Lopez-Fernandez MF, Rodriguez-Trillo A, Loures E, Cid AR, Bonanad S, Cabrera N, Moret A, Parra R, Mingot-Castellano ME, Balda I, Altisent C, Perez-Montes R, Fisac RM, Iruin G, Herrero S, Soto I, de Rueda B, Jimenez-Yuste V, Alonso N, Vilarino D, Arija O, Campos R, Paloma MJ, Bermejo N, Toll T, Mateo J, Arribalzaga K, Marco P, Palomo A, Sarmiento L, Inigo B, Nieto MM, Vidal R, Martinez MP, Aguinaco R, Cesar JM, Ferreira M, Garcia-Frade J, Rodriguez-Huerta AM, Cuesta J, Rodriguez-Gonzalez R, Garcia-Candel F, Cornudella R, Aguilar C, Borrás 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; **115**: 40-50.
22. Fidalgo T, Salvado R, Corrales I, Pinto SC, Borrás N, Oliveira A, Martinho P, Ferreira G, Almeida H, Oliveira C, Marques D, Goncalves E, Diniz M, Antunes M, Tavares A, Caetano G, Kjollerstrom 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; **116**: 17-31.
23. Cabrera N, Casana P, Cid AR, Moret A, Moreno M, Palomo A, Aznar JA. First application of MLPA method in severe von Willebrand disease. Confirmation of a new large VWF gene deletion and identification of heterozygous carriers. *Br J Haematol* 2011; **152**: 240-2.
  24. Aquila M, Bottini F, DI Duca M, Vijzelaar R, Molinari AC, Bicocchi MP. Multiplex ligation-dependent probe amplification to detect a large deletion within the von Willebrand gene. *Haemophilia* 2009; **15**: 1346-8.
  25. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, Lyon E, Ward BE. ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. *Genet Med* 2008; **10**: 294-300.
  26. Wallis Y, Payne S, McAnulty C, Bodmer D, Sistermans E, Robertson K, Moore D, Abbs S, Deans Z and Devereau A. ACGS & VKGL Practice Guidelines for the Evaluation of Pathogenicity and the Reporting of Sequence Variants in Clinical Molecular Genetics, 2013. Available at [http://www.acgs.uk.com/media/774853/evaluation\\_and\\_reporting\\_of\\_sequence\\_variants\\_bpgs\\_june\\_2013\\_-\\_finalpdf.pdf](http://www.acgs.uk.com/media/774853/evaluation_and_reporting_of_sequence_variants_bpgs_june_2013_-_finalpdf.pdf). Accessed January 10, 2017
  27. Goodeve A, Eikenboom J, Castaman G, Rodeghiero F, Federici AB, Batlle J, Meyer D, Mazurier C, Goudemand J, Schneppenheim R, Budde U, Ingerslev J, Habart D, Vorlova Z, Holmberg L, Lethagen S, Pasi J, Hill F, Hashemi SM, Baronciani L, Hallden C, Guilliat A, Lester W, Peake I. 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 (MCMMDM-1VWD). *Blood* 2007; **109**: 112-21.
  28. James PD, Notley C, Hegadom C, Leggo J, Tuttle A, Tinlin S, Brown C, Andrews C, Labelle A, Chirinian Y, O'Brien L, Othman M, Rivard G, Rapson D, Hough C, Lillicrap D. The mutational spectrum of type 1 von Willebrand disease: results from a Canadian cohort study. *Blood* 2007; **109**: 145-54.
  29. Standen GR, Bignell P, Bowen DJ, Peake IR, Bloom AL. Family studies in von Willebrand's disease by analysis of restriction fragment length polymorphisms and an intragenic variable number tandem repeat (VNTR) sequence. *Br J Haematol* 1990; **76**: 242-9.
  30. Othman M, Notley C, Lavender FL, White H, Byrne CD, Lillicrap D, O'Shaughnessy DF. Identification and functional characterization 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**: 4330-6.
  31. Weiss HJ, Meyer D, Rabinowitz R, Piétu G, Girma J, Vivic WJ, Rogers J. Pseudo-von Willebrand's disease: an intrinsic platelet defect with aggregation by unmodified human factor VIII/von Willebrand factor and enhanced adsorption of its high molecular weight multimers. *N Engl J Med* 1982; **306**: 326-33.

32. Ribba AS, Loisel I, Lavergne JM, Juhan-Vague I, Obert B, Cherel G, Meyer D, Girma JP. Ser968Thr mutation within the A3 domain of von Willebrand factor (VWF) in two related patients leads to a defective binding of VWF to collagen. *Thromb Haemost* 2001; **86**: 848-54.
33. Riddell AF, Gomez K, Millar CM, Mellars G, Gill S, Brown SA, Sutherland M, Laffan MA, McKinnon TA. Characterization of W1745C and S1783A: 2 novel mutations causing defective collagen binding in the A3 domain of von Willebrand factor. *Blood* 2009; **114**: 3489-96.
34. Flood VH, Schlauderaff AC, Haberichter SL, Slobodianuk TL, Jacobi PM, Bellissimo DB, Christopherson PA, Friedman KD, Gill JC, Hoffmann RG, Montgomery RR. Crucial role for the VWF A1 domain in binding to type IV collagen. *Blood* 2015; **125**: 2297-304.
35. Lyons SE, Cooney KA, Bockenstedt P, Ginsburg D. Characterization of Leu777Pro and Ile865Thr type IIA von Willebrand disease mutations. *Blood* 1994; **83**: 1551-7.
36. 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). *J Clin Invest* 1986; **77**: 947-51.
37. Schneppenheim R, Michiels JJ, Obser T, Oyen F, Pieconka A, Schneppenheim S, Will K, Zieger B, Budde U. A cluster of mutations in the D3 domain of von Willebrand factor correlates with a distinct subgroup of von Willebrand disease: type 2A/IIIE. *Blood* 2010; **115**: 4894-901.
38. Veyradier A, Boisseau P, Fressinaud E, Caron C, Ternisien C, Giraud M, Zawadzki C, Trossaert M, Itzhar-Baikian N, Dreyfus M, d'Oiron R, Borel-Derlon A, Susen S, Bezieau S, Denis CV, Goudemand J. A Laboratory Phenotype/Genotype Correlation of 1167 French Patients From 670 Families With von Willebrand Disease: A New Epidemiologic Picture. *Medicine (Baltimore)* 2016; **95**: e3038
39. Schneppenheim R, Brassard J, Krey S, Budde U, Kunicki TJ, Holmberg L, Ware J, Ruggeri ZM. Defective dimerization of the von Willebrand factor subunits due to a Cys->Arg mutation in type IID von Willebrand disease. *Proc Natl Acad Sci USA* 1996; **93**: 3581-6.
40. Holmberg L, Karpman D, Isaksson C, Kristoffersson A-C, Lethagen S, Schneppenheim R. Ins405AsnPro mutation in the von Willebrand factor propeptide in recessive type 2A (IIC) von Willebrand's disease. *Thromb Haemost* 1998; **79**: 718-22.

Table 1. Splice Site Prediction. Software tools available to assess the likely impact of a mutation on mRNA splicing.

<b>Program name</b>	<b>Web site address</b>
Splice Prediction Module	<a href="http://www.interactive-biosoftware.com/doc/alamut-visual/2.6/splicing.html">http://www.interactive-biosoftware.com/doc/alamut-visual/2.6/splicing.html</a>
MaxEntScan (5')	<a href="http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html">http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html</a>
MaxEntScan (3')	<a href="http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html">http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html</a>
NNSplice (FruitFly)	<a href="http://www.fruitfly.org/seq_tools/splice.html">http://www.fruitfly.org/seq_tools/splice.html</a>
GeneSplicer	<a href="http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml">http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml</a>
Human Splicing Finder	<a href="http://www.umd.be/HSF3/">http://www.umd.be/HSF3/</a>
NetGene 2	<a href="http://www.cbs.dtu.dk/services/NetGene2/">http://www.cbs.dtu.dk/services/NetGene2/</a>
MIT Splice Predictor	<a href="http://genes.mit.edu/GENSCAN.html">http://genes.mit.edu/GENSCAN.html</a>
Alternative Splice Site Predictor	<a href="http://wangcomputing.com/assp/">http://wangcomputing.com/assp/</a>

Use at least 3 of the reported predictions tools to reach a consensus result. Three of these prediction programs should agree, with a minimum reduction of 10% in the efficiency of a splice site being necessary to consider a deleterious effect on splicing.





Table 2. Amino acid prediction. Software tools available to assess the potential damaging effect on protein structure or function of a single nucleotide variant. The following programs are grouped dependent on type of tool. A minimum of about five tools, that rely on different types of algorithm, should be used.

<b>Program name</b>	<b>Type of tool</b>	<b>Web site address</b>
Domain Context	Evolutionary Conservation	<a href="http://pfam.xfam.org/">http://pfam.xfam.org/</a>
Domain Context	Evolutionary Conservation	<a href="http://www.uniprot.org/">http://www.uniprot.org/</a>
SNPs&GO	Machine-learning (SVM)	<a href="http://snps.biofold.org/snps-and-go/">http://snps.biofold.org/snps-and-go/</a>
MutPred	Supervised-learning	<a href="http://mutpred.mutdb.org/">http://mutpred.mutdb.org/</a>
PMut	Supervised-learning	<a href="http://mmb.pcb.ub.es/pmut2017/analyses/new/">http://mmb.pcb.ub.es/pmut2017/analyses/new/</a>
HANSA	Supervised-learning	<a href="http://www.cdfd.org.in/HANSA/">http://www.cdfd.org.in/HANSA/</a>
SNAP	Supervised-learning	<a href="https://roslab.org/services/snap/">https://roslab.org/services/snap/</a>
Align GVGD	Sequence/evolutionary conservation	<a href="http://agvgd.hci.utah.edu/">http://agvgd.hci.utah.edu/</a>
Panther Classification System Analysis	Sequence/evolutionary conservation	<a href="http://www.pantherdb.org/tools/csnpScoreForm.jsp">http://www.pantherdb.org/tools/csnpScoreForm.jsp</a>
PROVEAN	Sequence/evolutionary conservation	<a href="http://provean.jcvi.org/genome_submit_2.php?species=human">http://provean.jcvi.org/genome_submit_2.php?species=human</a>
SIFT Analysis	Sequence/evolutionary conservation	<a href="http://sift.bii.a-star.edu.sg/">http://sift.bii.a-star.edu.sg/</a>
MutationAssessor	Sequence/evolutionary conservation	<a href="http://mutationassessor.org/">http://mutationassessor.org/</a>
PolyPhen2 Analysis	Protein sequence and structure based	<a href="http://genetics.bwh.harvard.edu/pph2/">http://genetics.bwh.harvard.edu/pph2/</a>
SNPs 3D Analysis	Protein sequence and structure based	<a href="http://www.snps3d.org/">http://www.snps3d.org/</a>
Mutation Taster	Protein sequence and structure based	<a href="http://www.mutationtaster.org/">http://www.mutationtaster.org/</a>
Effect according to Russell	Severity of amino Acid Substitution	<a href="http://www.russelllab.org/aas/">http://www.russelllab.org/aas/</a>

Table 3. A strategy to identify the molecular defects in biochemical characterized type 2 VWD patients using PCR and Sanger sequencing.

VWD type	2A(IIA)	2A(II E)	2A*(II C)	2A(II D)	2B	2M	2M(CB)	2N*
Domain	A2	D3	D1-D2	CK	A1	A1	A3	D'-D3
Exon	3' portion of 28	22, 25-27 and 5' portion of 28	2-17	51-52	5' portion of 28	5' portion of 28	29-32	17-25

Each VWD type 2 variant results from sequence changes localized in a specific VWF functional domain. Although, some exceptions to this model have been reported, the large majority of the variants identified fit well with this pattern. Recessive variants (\*) are present in either homozygous or compound heterozygous form. In the latter it is likely that only the missense variant is responsible for patient type 2 phenotype, whereas the second defect may result in a null allele. Only missense variant may be identified using this approach and the recognition of the second variant will require evaluation of the entire *VWF* coding sequence.