INTRODUCTION

Von Willebrand disease (VWD) is the most common hereditary bleeding disorders that occurs due to deficiency or abnormality of von Willebrand Factor (VWF). VWF is an elongated multimeric plasma glycoprotein that plays essential roles in primary and secondary haemostasis. The gene that encodes VWF (VWF) is positioned on the short arm of chromosome 12 (12p 13.2) and comprises 52 exons, spanning~178 kb of genomic sequence. Furthermore, a partial pseudogene with high homology (97%) to exons 23-34 of the VWF is
localized on chromosome 22. The pro-pre-VWF protein is composed of a 741-amino acid (aa) propeptide (D1-D2 domains), and a highly glycosylated 2050-aa mature VWF subunit (domains of D’-D3-A1-A2-A3-D4-C1-C2-C3-C4-C5-C6-C7-K). The D domains each comprise a VW domain, a C8 fold, a trypsin inhibitor-like (TIL) structure and an E module. The exceptions are where D’ lacks the VW domain and C8 fold, and the D4 domain lacks the E module but encompasses subdomain D4N. The C-terminal domains as well as D1, D2, D’ and D3 domains are characterized as structural domains which are required for dimerization and multimerization process, respectively. The designated functional domains include A1-A3 and D’-D3 domains.

von Willebrand disease is categorized into quantitative types 1 and 3 or qualitative type 2. Type 3 VWD, the most severe form of VWD, is defined as the virtual absence of VWF (VWF:Ag < 5 IU/dL). It is inherited as an autosomal recessive trait due to either homozygous or compound heterozygous of VWF mutations. It affects 0.1 to 0.3 per million people, with an increased incidence in regions with a high rate of consanguinity. In general, type 3 VWD patients present moderate to severe bleeding symptoms. They exhibit not only classic mucocutaneous bleeding (epistaxis, gum bleeding, easy bruising and menorrhagia), the most common complaints in VWD patients, but also hemarthroses and hematomas similar to that of moderate haemophilia. Nonetheless, most common complaints in VWD patients, but also hemarthroses and bleeding (epistaxis, gum bleeding, easy bruising and menorrhagia), the severe bleeding symptoms. They exhibit not only classic mucocutaneous bleeding (epistaxis, gum bleeding, easy bruising and menorrhagia), the most common complaints in VWD patients, but also hemarthroses and hematomas similar to that of moderate haemophilia. Nonetheless, they demonstrate variability in severity and type of bleeding episodes along with heterogeneity in molecular aetiology which is indicated with a broad spectrum of identified VWF mutations.

In the current study, we aimed to characterize the genotype and clinical bleeding of a large cohort of type 3 VWD patients of Pakistani population for the first time to designate the phenotype-genotype correlations. This study will elucidate the distinct picture of the genetic aetiology of type 3 VWD in Pakistani population which help in accurate diagnosis, estimating bleeding risk for the type 3 VWD patients and eventually providing best therapy regimes.

2 | MATERIAL AND METHODS

2.1 | Patients

Forty-eight index patients (IPs) diagnosed with type 3 VWD from apparently unrelated families of various regions of Pakistan were registered in this study. The provisional diagnosis of type 3 VWD was made based on bleeding history and laboratory evaluations in respective collaborating centres based on ISTH-SSC VWF guidelines. All patients participating in this study gave written informed consent in local language, and study was approved by the Review Ethical Board of NIBD, Pakistan.

Peripheral blood samples from IPs were collected in tubes containing 3.2% sodium citrate and EDTA for coagulation tests and DNA extraction, respectively.

2.2 | Phenotypic evaluation

Coagulation studies: To confirm the type 3 VWD diagnosis, the conventional coagulation assays including VWF antigen levels (VWF:Ag) levels and factor VIII coagulant activity (FVIII:C) were performed as described somewhere else. Inhibitor assay against VWF: Screening for alloantibodies to VWF was performed using an ELISA-based test plus an inhibitory assay evaluating VWF-binding function to collagen (VWF:C).

Bleeding history: A validated questionnaire was administered to our cohort of patients to record detailed bleeding episodes of patients.

2.3 | VWF mutation analysis

DNA extraction: Genomic DNA was isolated from EDTA blood using a QiAamp DNA Blood Mini Kit (Qiagen®, Germany). DNA sequencing and mutation detection: Essential regions of VWF were amplified by exon-specific polymerase chain reaction (PCR) using primer sets which covered exons 2-52 and exon/intron boundaries. Primer sequences and PCR reaction conditions are offered upon request. The amplified products of VWF-coding regions were subsequently subjected to Sanger direct sequencing using the Big-Dye Terminator Cycle Sequencing Kit v3.1 and the automated ABI 3500xl Genetic Analyzer (Applied Biosystems, CA, USA). Generated sequences were analysed using SeqScape software (Applied Biosystems, CA, USA) to detect candidate mutations.

Any detected VWF sequence variation was documented, and it was checked whether they had been earlier recorded in professional databases and published literature. They were sought in PubMed and the Online Mendelian Inheritance in Man (OMIM) (www.ncbi.nlm.nih.gov/omim), the Coagulation Factor Variant Databases portal supported by European Association for Haemophilia and Allied Disorders (EAHAD) (https://grenada.lumc.nl/LOVD2/VWF/home.php?select_db=vVWF), VWFdb Hemobase (http://www.vwf.hemobase.com), Human Gene Mutation Database (HGMD) (http://www.hgmd.cf.ac.uk/ac/index.php?gene=VWF), SNPdb (through NCBI; http://www.ncbi.nlm.nih.gov/SNP), all accessed March 2019.

In silico prediction: Pathogenicity of novel missense variants were evaluated by MutationTaster (http://www.mutationtaster.org/) which analyses several elements including evolutionary conservation, impact on protein features (e.g. protein folding, post-translational modifications, and protein binding sites) and effect on splice sites. Additionally, other prediction bioinformatics tools were used to characterize the impact of the novel missense changes on VWF protein structure and function, they are ConSurf server (http://consurf.tau.ac.il/2016/) Polymorphism Phenotyping v2 (Polyphen-2; http://genetics.bwh.harvard.edu/pph2/), Provean (http://provean.jcvi.org/index.php/), Sorting Intolerant From Tolerant, SIFT (http://sift.jcvi.org/) and Mupro (http://mupro.proteomics.ics.uci.edu/), all accessed May 2018.

Pathogenicity of potential splice site mutations was evaluated by splice site prediction tools Human splicing finder (HSF) version 3.1 (http://www.umd.be/HSF3/HSF.shtml) and WebGene-SpliceView (http://bioinfo.itb.cnr.it/~webgene/wwwspliceview.html), both accessed April 2017.
2.4 | In silico structural analysis

To evaluate the putative structural impact of the novel missense variants, we performed in silico structural analysis. We downloaded the crystal structures of A3 (PDB ID: 4dmu; resolution: 2.8 Å) and CTCK (PDB ID: 4nt5; resolution: 3.28 Å) domains from the structure database, while other domains like C1 and D2 domains which do not have any biophysical structure were modelled on the threading modelling server ITASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/, accessed November 2018). The models were generated by submitting the sequence based on the currently accepted domain annotation into the server. The C1 domain sequence was submitted with additional restraints, that is the PDB ID 1u5m as a template and known disulphide cysteines interatomic distances. The D domain was submitted only with known disulphide cysteines interatomic distances as restraint. The local molecular environment of each reported residue was inspected on the structures/models to understand their impact on the structure and function of the protein. All structural visualization, analysis and image rendering were performed on the YASARA platform.

3 | RESULTS

3.1 | Patients: Phenotype and bleeding episodes

A total of 48 IPs diagnosed with type 3 VWD, all coming from consanguineous marriages, were included in this study. From 48 patients, 26 (54%) were males, and 22 (46%) were females. The mean age was 11.5 ± 4.1 (ranging from 6-22). Easy bruising was the most common bleeding manifestation (88.8% of patients) followed by bleeding from minor wounds (62.5%) and nose-bleeding (42.2%). Patients also exhibited gum bleeding (28.9%), post-dental extraction bleeding (26.7%), muscle haematomas (17.8%) and joint bleeding (13.3). Furthermore, two of the patients (IP-45 and IP-48) exhibited central nervous bleeding. Besides, 6 out of 22 females showed menorrhagia (Table 1). All the patients received on-demand therapy in the form of fresh-frozen plasma and/or cryoprecipitate. Few of them were given FVIII/VWF concentrate along with cryoprecipitate.

VWF:Ag levels were less than or equal to 3% in all patients, and FVIII:C levels were ranging from <1 to 6.6 IU/dL (Table 1). In none of the patients, alloantibodies against VWF was detected.

3.2 | Genotype

We identified putative VWF mutations in 46 out of 48 IPs (95.8%). Forty-four of the IPs were homozygous for the given detected gene variations, one patient was compound heterozygous, and one patient has more than two variants (two sequence variations as homozygous and one heterozygous variation, IP-16). In general, ninety-three putative mutations were detected that the majority of them were null. The gene variations comprise 27 missense mutations (~30%), 23 nonsense mutations (~25%), 22 gene conversions (~23%), 10 splice site alterations (~11%), 5 small insertions (~5%), 4 small deletions (~4%) and 2 large deletion (~2%) (Figure 1A). Many of the variants were detected more than one time, either as homozygosity status or their incidence in more than one patient. In total, there are therefore twenty-nine distinctive gene variations in this cohort (66% null mutations and 34% missense variations) from which 10 (34.5%) are reported for the first time.

Missense Mutation: Of the 48 IPs, 14 cases harbour missense variations which are either as homozygous or as compound heterozygous (combined with null alleles). A total of 27 missense variations (including nine different missense variations) were identified in this cohort of type 3 VWD patients. Three of the missense mutations (p.Cys633Arg, p.Phe1369Ile, p.Leu1781Trp) were reported in more than one patient (Table 1). Six of missense substitutions were reported for the first time. The novel variants were p.Cys608Trp, p.Val1485Leu, p.Leu1781Trp,
| Patient ID | Sex/Age (Year) | VWF:Ag IU/dL | FVIII:C IU/dL | Bleeding Symptoms | Mutation Type | Nt Change | aa-Change | Domain Subdomain | Zygosity |
|------------|----------------|---------------|---------------|-------------------|---------------|-----------|-----------|-----------------|----------|
| IP-01      | M/6 ≤3         | 1.1           | 1.2,3,8       | Nonsense          | c.100C > T    | p.Arg34*  | D1/VWD1   | Hmz             |
| IP-02      | F/9 <3         | 1             | 2.4           | Missense          | c.257T > A    | p.Val86Glu| D1/VWD1   | Hmz             |
| IP-03      | F/9 ≤3         | 1.7           | 2.3           | Insertion         | c.780-781insG | p.Leu261Alafs*41| D1/C8-1 | Hmz             |
| IP-04      | M/15 <3        | 1.4           | 1.2,3         | Insertion         | c.780-781insG | p.Leu261Alafs*41| D1/C8-1 | Hetz            |
| IP-05      | M/7 <3         | 2.2           | 1.2           | Small Deletion    | c.871delT²    | p.Cys291Alafs*166 | D1/C8-1 | Hmz             |
| IP-06      | M/7 ≤3         | 1.3           | 1.2,3,4,5     | Nonsense          | c.970C > T    | p.Arg324* | D1/TIL1   | Hmz             |
| IP-07      | F/12 <3        | 1.1           | No            | Nonsense          | c.970C > T    | p.Arg324* | D1/TIL1   | Hmz             |
| IP-08      | M/8 ≤3         | 1.7           | 2,3,4,5,8     | Nonsense          | c.1117T > C   | p.Arg373* | D1-E1     | Hmz             |
| IP-09      | F/14 ≤3        | 1.2           | 2             | Nonsense          | c.1117T > C   | p.Arg373* | D1-E1     | Hmz             |
| IP-10      | F/21 <3        | 3.7           | _             | Splice Site       | c.7820 + 3A>C | D2/C8-2 | Hmz       |
| IP-11      | F/6 ≤3         | 1.1           | 2.8           | Splice Site       | c.1729 + 3A>C | D2/C8-2 | Hmz       |
| IP-12      | M/8 ≤3         | 1.4           | 1.4           | Missense          | c.1824C > G   | p.Cys608Trp| D2/C8-2 | Hmz             |
| IP-13      | F/8 ≤3         | 1.7           | 2             | Missense          | c.1897T > C   | p.Cys633Arg| D2/C8-2 | Hmz             |
| IP-14      | M/15 ≤3        | 1.33          | 2             | Missense          | c.1897T > C   | p.Cys633Arg| D2/C8-2 | Hmz             |
| IP-15      | M/12 ≤3        | 1             | 1,2,4,5       | Splice Site       | c.1729 + 3G>C | D3/VWD3 | Hmz       |
| IP-16      | F/12 ≤3        | 1.1           | 1.2,3         | Splice Site       | c.1729 + 3G>C | D3/VWD3 | Hmz       |
| IP-17      | F/7 ≤3         | 1.6           | 2             | Nonsense          | c.3030C > A   | p.Cys1101*| D3/C8-3 | Hmz             |
| IP-18      | F/11 ≤3        | 1.1           | 2             | Nonsense          | c.3030C > A   | p.Cys1101*| D3/C8-3 | Hmz             |
| IP-19      | M/11 ≤3        | 1.1           | 1,2,3,4,5     | Gene Conversion   | c.3797C > T   | p.Pro1266Leu| A1       | Hmz             |
|            |                |               |               |                   | c.3835G > A   | p.Val1279Ile| A1       |                 |
|            |                |               |               |                   | c.3931C > T   | p.Gln1311* | A1       |                 |
| IP-20      | M/14 ≤3        | 2.0           | 2,3,5         | Gene Conversion   | c.3797C > T   | p.Pro1266Leu| A1       | Hmz             |
|            |                |               |               |                   | c.3835G > A   | p.Val1279Ile| A1       |                 |
|            |                |               |               |                   | c.3931C > T   | p.Gln1311* | A1       |                 |
| IP-21      | M/9 ≤3         | 1.3           | 3,5,6         | Gene Conversion   | c.3797C > T   | p.Pro1266Leu| A1       | Hmz             |
|            |                |               |               |                   | c.3835G > A   | p.Val1279Ile| A1       |                 |
|            |                |               |               |                   | c.3931C > T   | p.Gln1311* | A1       |                 |
| IP-22      | F/7 ≤3         | 1             | 2,8           | Gene Conversion   | c.3835G > A   | p.Val1279Ile| A1       | Hmz             |
|            |                |               |               |                   | c.3931C > T   | p.Gln1311* | A1       |                 |
|            |                |               |               |                   | c.4027A > G   | p.Lle1343Val| A1       |                 |
|            |                |               |               |                   | c.4105T > A   | p.Phe1369Ile| A1       |                 |
| IP-23      | M/7 ≤3         | 2.5           | 1,2,3         | Gene Conversion   | c.3931C > T   | p.Gln1311* | A1       | Hmz             |
|            |                |               |               |                   | c.4027A > G   | p.Lle1343Val| A1       |                 |
|            |                |               |               |                   | c.4105T > A   | p.Phe1369Ile| A1       |                 |
| IP-24      | M/9 ≤3         | 1.5           | 2,3,4,5,6,8   | Gene Conversion   | c.3931C > T   | p.Gln1311* | A1       | Hmz             |
|            |                |               |               |                   | c.4027A > G   | p.Lle1343Val| A1       |                 |
|            |                |               |               |                   | c.4105T > A   | p.Phe1369Ile| A1       |                 |
| IP-25      | M/11 ≤3        | 2.0           | 1,2,3,4,5,6   | Gene Conversion   | c.3931C > T   | p.Gln1311* | A1       | Hmz             |
|            |                |               |               |                   | c.4027A > G   | p.Lle1343Val| A1       |                 |
|            |                |               |               |                   | c.4105T > A   | p.Phe1369Ile| A1       |                 |

(Continues)
The potential impact of these gene variations was investigated using several in silico programs; the predicted results are presented in Table 2. The in silico MuPro program evaluated the effect of the newly reported missense variants on protein stability by calculating the value of energy change. It predicted that all of these novel missense substitutions decrease protein stability (Table 2). On the other hand, other four in silico tools (MutationTaster, PolyPhen-2, SIFT and ConSurf), which predict the pathogenicity of the missense variant by evaluating either their impact on protein features and/or calculating positional conservation using multiple alignments, were applied. Of six novel missense variants, three of them (p.Cys608Trp, p.Val2326Met and p.Cys2739Ser) were predicted to be damaging by all of these prediction programs (Table 2, dark grey rows). The novel missense

| Patient ID | Sex/Age (Year) | VWF:Ag IU/dL | FVIII:C IU/dL | Bleeding Symptoms | Mutation Type | Nt Change | aa-Change | Domain Subdomain | Zygosity |
|------------|----------------|--------------|--------------|------------------|---------------|-----------|-----------|------------------|----------|
| IP-26      | F/11 ≤3        | <1           | 1,2,3        | Gene Conversion | c.3931C > T   | p.Gln1311* | A1        | Hmz              |
| IP-27      | M/12 <3        | 2.2          | 2.3          | Gene Conversion | c.3931C > T   | p.Gln1311* | A1        | Hmz              |
| IP-28      | F/17 <3        | 1.5          | 1,2,3        | Gene Conversion | c.3931C > T   | p.Gln1311* | A1        | Hmz              |
| IP-29      | M/12 ≤3        | 1.1          | 2,3,4        | Gene Conversion | c.3931C > T   | p.Gln1311* | A1        | Hmz              |
| IP-30      | F/7 ≤3         | <1           | 2            | Nonsense        | c.3931C > T   | p.Gln1311* | A1        | Hmz              |
| IP-31      | M/12 ≤3        | <1           | 2,3,5        | Missense       | c.4027A > G   | p.Ile1343Val | A1        | Hmz              |
| IP-32      | F/15 ≤3        | 1.1          | 2            | Missense       | c.4027A > G   | p.Ile1343Val | A1        | Hmz              |
| IP-33      | M/11 ≤3        | 1.4          | 1,2,3,6      | Small Deletion | c.4176delG   | p.Arg1392fs*12 | A1        | Hmz              |
| IP-34      | M/9 ≤3         | <1           | 2,3,5        | Insertion      | c.4413-4414insCC | p.Asp1472Profs*53 | A1        | Hmz              |
| IP-35      | M/15 ≤3        | 2.5          | 2            | Nonsense       | c.4975C > T   | p.Arg1659* | A2        | Hmz              |
| IP-36      | F/18 <3        | 2.0          | 2,7          | Nonsense       | c.4975C > T   | p.Arg1659* | A2        | Hmz              |
| IP-37      | M/10 ≤3        | 1.2          | 1,2,3        | Missense       | c.5342T > G   | p.Leu1781Trp | A3        | Hmz              |
| IP-38      | F/11 ≤3        | 2.4          | 2,7          | Missense       | c.5342T > G   | p.Leu1781Trp | A3        | Hmz              |
| IP-39      | F/14 ≤3        | 2.5          | 2,7          | Missense       | c.5342T > G   | p.Leu1781Trp | A3        | Hmz              |
| IP-40      | F/17 <3        | 3.0          | _            | Nonsense       | c.5557C > T   | p.Arg1853* | A3        | Hmz              |
| IP-41      | M/6 ≤3         | 2.6          | 1,2,4,5      | Missense       | c.6709T > C   | p.Cys2237Arg | D4/TIL4   | Hmz              |
| IP-42      | M/8 ≤3         | 2.9          | 1,2,3,4,5    | Missense       | c.6976G > A   | p.Val2326Met | C1        | Hmz              |
| IP-43      | F/22 <3        | 2.7          | _            | Missense       | c.6982G > A   | p.Val2326Asn | C1        | Hmz              |
| IP-44      | M/18 <3        | 2.5          | 1,2,3,6      | Splice Site    | c.8155 + 6T > A | p.As2328Asn | C1        | Hmz              |
| IP-45      | F/10 <3        | 1.9          | 7,8,9        | Missense       | c.8216G > C   | p.Cys2739Ser | CK        | Hmz              |
| IP-46      | M/19 ≤3        | 1.4          | 1,2,3,8      | Large Deletion | p.Cys2739Ser | CK        | Hmz      |                  |
| IP-47      | M/14 <3        | 6.6          | 1,2,3,4,6,8  | No Mutation    |               |            |          |                  |
| IP-48      | F/15 ≤3        | 1.1          | 2,3,4,7,9    | No Mutation    |               |            |          |                  |

Note: The numbers in bleeding symptoms column represent: 1 = Epistaxis, 2 = easy bruising, 3 = bleeding from minor wounds, 4 = bleeding from oral cavity, 5 = bleeding after tooth extraction, 6 = joint bleeding, 7 = menorrhagia, 8 = muscle haematomas, 9 = CNS bleeding. No bleeding history was available for the IP-40 and IP-43 (marked with _). Abbreviations: aa, amino acid; BS, bleeding score; FVIII:C, FVIII coagulant activity; Hetz, heterozygous; Hmz, homozygous; Nt, nucleotide; VWF:Ag, VWF antigen levels.

*Indicates novel mutations.
*Symbolizes the translation stop codon.
### TABLE 2  Prediction of the applied bioinformatics programs (MutationTaster, PolyPhen-2, SIFT, Provean, and MUpro) on the pathogenicity of novel missense variations

| #  | Nt. Change | aa Exchange | Domain | Mutation Taster Prediction (Evolutionary conservation / Protein features / Splice site) | ConSurf Prediction (Conservation Score) | PolyPhen-2 Prediction (Probability Score) | SIFT Prediction (SIFT Score) | MUpro Prediction (Protein Structure Stability (delta delta G)) |
|----|------------|-------------|--------|----------------------------------------------------------------------------------|----------------------------------------|------------------------------------------|-------------------------------|----------------------------------------------------------|
| 1  | c.1824C>G  | p.Cys608Trp | D2     | Probably deleterious (Conserved/No/No)                                            | Conserved (9)                          | Probably Damaging (0.999)                | Damaging (0)                  | Decrease Stability (-1.5181447)                           |
| 2  | c.4453G>C  | p.Val1485Leu| A2     | Probably harmless (Tolerated/Protein features might be affected/Acc gained)       | Variable (1)                          | Benign (0.012)                          | Tolerated (0.28)              | Decrease Stability (-0.31612469)                          |
| 3  | c.5342T>G  | p.Leu1781Trp| A3     | Probably harmless (Partly conserved/Protein features might be affected/May affect splicing) | Variable (1)                          | Probably Damaging (0.945)               | Damaging (0.01)               | Decrease Stability (-1.5005777)                           |
| 4  | c.6976G>A  | p.Val2326Met| C1     | Probably deleterious (Conserved/Protein features might be affected/Donor lost)     | Conserved (7)                          | Probably Damaging (1.000)               | Damaging (0.04)               | Decrease Stability (-0.43462811)                          |
| 5  | c.6982G>A  | p.Asp2328Asn| C1     | Probably deleterious (Conserved/Protein features might be affected/May affect splicing) | Conserved (9)                          | Probably Damaging (1.000)               | Tolerated (0.12)              | Decrease Stability (-0.82981467)                          |
| 6  | c.8216G>C  | p.Cys2739Ser| CK     | Probably deleterious (Conserved/Protein features might be affected/No)            | Conserved (9)                          | Probably Damaging (0.999)               | Damaging (0.01)               | Decrease Stability (-1.4068571)                           |

The rows of table are colored in a progressive gradient of gray. The darker the shade of gray is the more number of in silico tools predict deleteriousness of the mutant the row represents; for example the variants p.Cys608Trp, p.Val2326Met and p.Cys2739Ser in dark gray rows were predicted to be damaging by all the in silico prediction tools. MutationTaster predicts an alteration as disease-causing (i.e. probably deleterious) or polymorphism (i.e. probably harmless). Moreover, it provides a summary listing the most prominent features of the analysed alteration including evolutionary conservation, impact on protein features and their probable effect on splice sites. PolyPhen-2, SIFT, and Consurf predict the possible impact of the amino acid (aa) substitutions on the function of the protein using multiple alignment pipeline and building conservation profiles. The output of the PolyPhen-2 is a prediction of probably damaging, possibly damaging, or benign, accompanied by a numerical score ranging from 0.0 (benign) to 1.0 (damaging). SIFT predicts the variants as tolerated or deleterious substitutions, along with a normalized probability score; scores < 0.05 are predicted to be deleterious and those ≥ 0.05 are predicted to be tolerated. ConSurf provides a conservation score ranging from 1 to 9 indicating the conservation degree of the corresponding amino acids at mutated residues; conservation scale is nominated as conserved (scores 7, 8 and 9), average (scores 4, 5 and 6) and variable (scores 1, 2 and 3). MUpro is a set of machine learning programs to predict the impact of the sequence variations on protein stability. Here the prediction of the value of energy change (delta delta G) is provided. A score < 0 means the substitution decreases the protein stability, and a score > 0 implies the substitution increases the protein stability.
variations p.Leu1781Trp and p.Asp2328Asn were predicted to be deleterious by at least two of the in silico prediction tools (but not by all of the programs) (Table 2, light grey rows). The missense substitution p.Val1485Leu was anticipated to be benign by all of the four in silico tools (Table 2, white row).

**Gene Conversion:** Gene conversions between VWF (on chromosome 12) and VWF pseudogene (on chromosome 22) were detected in 11 out of 48 IPs as homozygous in our cohort. The gene conversions are recognized by multiple point variations including missense and nonsense, and they are reported earlier with various length and composition. In three of IPs, the converted gene includes sequence variations p.Pro1266Leu, p.Val1279Ile and p.Gln1311*. In one of the patients, it comprises four variations p.Val1279Ile, p.Gln1311*, p.Ile1343Val and p.Phe1369Ile. Lastly, 8 of 11 patients bearing gene conversion have variants p.Gln1311*, p.Ile1343Val and p.Phe1369Ile. In one of the patients, it comprised four variations p.Val1279Ile, p.Gln1311*, p.Ile1343Val and p.Phe1369Ile (Table 1). Nevertheless, in all cases the nonsense mutation p.Gln1311* was shared which would result in the same truncated protein in all of these patients.

**Nonsense Mutations:** Twelve of 48 IPs have nonsense mutations, constituting ~25% of mutations. In total, seven different mutations were identified (p.R34*, p.R324*, p.R373*, p.C1101*, p.Q1311*, p.R1659* and p.R1853*) of which five were converting Arginine amino acid to stop codon. Mutations p.R324*, p.R373*, p.C1101* and p.R1659* were found in more than one patient (Table 1). All of the nonsense mutations were previously reported.

**Insertions/Deletions:** We identified two different insertion mutations with a gain of one or two nucleotides (c.780_781insG and c.4413_4414insCC). The novel small insertion c.780_781insG was located in a stretch of guanine residues, and the previously reported insertion c.4413_4414insCC situated on a stretch of cytosine residues. The single nucleotide insertion mutation (c.780_781insG) was detected in two patients of our cohort.

Furthermore, we detected two novel small deletions which were a single nucleotide deletion (c.871delT as homozygous and c.4176delG as compound heterozygous).

All of the small insertions and deletions discovered in our cohort of type 3 VWD patients cause the shift of the reading frame leading to a premature termination code and consequently non-functional protein product.

**Large Deletion:** A homozygous large deletion of exon 40 was detected in one of the patients by lack of PCR amplification product of exon 40 repetitively. However, the occurrence of this large deletion was not confirmed by further gene dosage-based analysis (e.g. multiplex ligation-dependent probe amplification, MLPA) due to an insufficient amount of DNA material from the index patient.

**Splice Site Mutation:** Three potential splice site alterations including one previously reported mutation c.1729 + 3G>A and two variants c.2820 + 5G>C and c.8155 + 6T>A were detected in this study. The variants c.1729 + 3G>A and c.2820 + 5G>C were identified in more than one IP (each in two patients), they were inherited either as

---

**FIGURE 2** In silico views of the molecular position of the novel mutated residues depicted on structural models of D2, A3, C1, C2 and CTCK domains. Panel A. illustrates mutated residue p.Cys608 on the threaded model of domain D2. The protein backbones in this and all subsequent panels except Panel B are depicted in ribbon format and coloured according to their secondary structure (Beta regions: Red, Short Helix: Green, Coiled regions: Cyan). All cysteines are depicted in stick format and coloured blue unless they are the residues on which the mutation has been reported. All mutated residues have been illustrated in grey coloured stick format. Panel B shows the position of variation p.Leu1781Trp on the crystal structure of domain A3. The protein backbone is represented in grey coloured tube format. Channels and voids calculated from the BetaCavityWeb server are depicted with red and green colours (http://voronoi.hanyang.ac.kr/betacavity web/). The location of the affected residue is indicated in this panel although the residue itself is not shown. Panel C demonstrates two affected residues p.Val2326 and p.Asp2328 on the threaded model of domain C1. Panel D exhibits the crystal structural domain CTCK (C-terminal cysteine knot) and position of mutated residue p.Cys2739.
homozygous or as compound heterozygous. The potential impact of the novel splice site variants on splicing was evaluated by using in silico tools HSF and WebGene-SpliceView. The analyses by both tools predicted that the substitution G > C at donor site c.2820 + 5G>C of exon 21 might abolish the donor splice site. On the other hand, the HSF and WebGene-SpliceView tools predicted that variant c.8155 + 6T>A at donor site of exon 50 may lead to the creation of a cryptic new acceptor or hindering the donor splice site, respectively. In total, the in silico analyses predicted that these two splice site variants might alter splicing accuracy and efficiency.

3.3 | In silico structural analysis

An inspection of the D2, C1 and CTKT domain residues on their respective models and structures allowed us to classify the proposed structural impact of the putative novel missense mutations reported in this study into two major types: a) variants of p.Val2326Met and p.Asp2328Asn which were close to cysteines in reduced form and b) variants p.Cys608Trp and p.Cys2739Ser which were on cysteines participating in domain stabilizing disulphide bonds. The reduced cysteines observed in these domains often participate in the interdomain or inter-subunit disulphide bonds and therefore critical to final VWF assembly. The cysteines already part of disulphide bonds contribute to domain stability or the ability of these domains to conformationally transition upon a change in pH. The net pathomolecular effect of the variants that occur on these cysteines or close to them is mediated through these cysteines. However, the extent of their effect differs based on the type of substitution. Additionally, the novel missense variation p.Leu1781Trp occurs on the A3 domain in a region that is closely packed. The A3 domain is a closely packed domain with a hydrophobic core with very few cavities, or voids is as evident from the red and green coloured regions (Figure 2B). The substitution of a comparatively small-sized Leu with a large hydrophobic Trp head would result in significant clashes locally and disturb the overall stability of the complete domain and protein.

The non-conservative substitution p.Val1485Leu lies on an approximately 80-amino acid long domain linker between domains A1 and A2 and is structurally uncharacterized. Since almost one-fourth of this linker is composed of either Valine or Leucine, this suggests that a Val to Leu substitution is unlikely to have a significant structural impact on it.

4 | DISCUSSION AND CONCLUSION

This study is the first report of detailed mutation spectrum in type 3 VWD in the Pakistani population with a high rate of consanguineous marriage. The cohort comprises 48 unrelated type 3 VWD patients, all coming from consanguinity marriage. In this study, we were able to detect both VWF mutations in ~96% of the patients (46 out of 48 IPs) which were mainly homozygous. The disease-causing variations in two remaining IPs might be large deletions, deep intronic variations or distant regulatory defects outside of VWF. DNA material was not sufficient to further do gene dosage-based analysis (MLPA) to detect any heterozygous large deletion in the later two IPs (without any detected gene variation) as well as in the patients with apparent homozygous mutations, if the failure to amplify deleted VWF allele is concealed by amplification of the second allele. Nonetheless, this probability is not high for our cohort due to the consanguinity.

In total, twenty-nine different gene defects were characterized in this cohort from which 10 (34%) are reported for the first time. Remarkably, previous cohort studies of type 3 VWD have persistently reported that a bulk of the putative mutations were reported for the first time.12,24,26-28 This highlights the high predisposition of VWF for sequence variations and mutational heterogeneity in type 3 VWD.

The candidate mutations were scattered throughout the entire VWF, not confined to the specific region or domain of VWF, as observed in the other cohorts.12,26,28,29 Consistent with the literature, the sequence variations leading to null alleles were predominant in our group of type 3 VWD (66%; including gene conversions, nonsense, splice site variations, small deletions, insertions and large deletion), while missense mutations comprise 34% of mutations. Nonetheless, the incidence rate of missense variants in our cohort (~34%) was slightly higher compared to the other type 3 VWD cohorts, which is commonly 18%-24%.12,16,26,27 This might be due to the high rate of consanguinity in our cohort, and a subsequently greater likelihood of homozygosity for missense mutations. The missense mutations were located in D1, D2, A1, A2, A3, D4, C1 and CK domains (Figure 1B). It can be expected that mutations in C-terminal domains (C1 and CK) and propeptide region (D1 and D2) may interfere with dimerization and polymerization of VWF monomers, respectively.30-32 On the other hand, the molecular mechanism of the mutations located in A domains (A1, A2, and A3) and D4 by which impair the total biosynthesis of VWF is remained to be understood by further gene expression analysis. Six of ten missense variations were novel. The pathogenicity of the novel missense variations was evaluated with prediction bioinformatics tools providing evolutionary conversation score as well as in silico structure analysis. Both analyses showed that five conservative missense variations p.Cys608Trp, p.Leu1781Trp, p.Val2326Met, p.Asp2328Asn and p.Cys2739Ser would have deleterious effects, and they were predicted to be disease-causing variants. We assume either they affect domain stability through interfering with the formation of intra-domain disulphide bonds (p.Cys608Trp, p.Cys2739Ser), or they have an impact on inter-domain disulphide bonds contributing to VWF assembly (p.Val2326Met and p.Asp2328Asn). On the other hand, variation p.Val1485Leu was assigned to be tolerable/probably harmless by all prediction tools; it was anticipated that substitution Val to Leu in this position has unlikely a significant structural impact. Nevertheless, the casualty of these novel missense variants cannot be confidently confirmed or excluded before further expression studies. Gene conversion mutations were identified in 11 of IPs in the current cohort (comprising ~23% of mutations). Correspondingly, gene conversions are repeatedly reported in the
neighbouring country population, India, this suggests conceivable founder origin of these mutations. However, the incidence of the gene conversion variations in our cohort (23%) was greater when is compared with those reported for different Indian cohorts, for example 18.5%, 10.2% and 4.5% reported by Gupta et al, Kasatkar et al and Elayaperumal et al, respectively.26,27,33

In conclusion, our study represents one of the few studies reporting the spectrum of mutations in type 3 VWD cohort with a high rate of consanguinity, providing further insights into the VWD pathogenesis in a consanguineous population.

DISCLOSURE

The authors stated that they had no interests which might be perceived as posing a conflict or bias.

AUTHOR CONTRIBUTION

S.A performed the experimental work and wrote the paper. H.Y designed the study, interpreted the data; wrote the paper. A.N recorded and evaluated the bleeding history and course of therapy of the patients. B.S performed in silico structural analysis. U.B evaluated the development of alloantibodies against VWF in the patients. A.P supervised the gene sequencing experiments. N.S, S.A, S.T, F.R, S.M and TS.S contributed to the patient's data. J.O supervised the study, reviewed and edited the manuscript.

ORCID

Hamideh Yadegari
https://orcid.org/0000-0002-2006-8497

Johannes Oldenburg
https://orcid.org/0000-0002-1585-4100

REFERENCES

1. Castaman G, Rodeghiero F, Tosetto A, et al. Hemorrhagic symptoms and bleeding risk in obligatory carriers of type 3 von Willebrand disease: An international, multicenter study. J Thromb Haemost. 2006;4(10):2164-2169.
2. James PD, Goode AC. von Willebrand disease. Genet Med. 2011;13(5):365-376.
3. Leebeek F, Eikenboom J. Von Willebrand's Disease. N Engl J Med. 2016;375(21):2067-2080.
4. Mancuso DJ, Tuley EA, Westfield LA, et al. Structure of the gene for human von Willebrand factor. J Biol Chem. 1989;264(33):19514-19527.
5. Goode AC. The genetic basis of von Willebrand disease. Blood Rev. 2010;24(3):123-134.
6. Gupta PK, Adamziki E, Budde U, et al. Gene conversions are a common cause of von Willebrand disease. Br J Haematol. 2005;130(5):752-758.
7. Zhou Y-F, Eng ET, Zhu J, Lu C, Walz T, Springer TA. Sequence and structure relationships within von Willebrand factor. Blood. 2012;120(2):449-458.
8. Lenting PJ, Casari C, Christophe OD, Denis CV. von Willebrand factor: the old, the new and the unknown. J Thromb Haemost. 2012;10(12):2428-2437.
9. Schnepfenreich M. The pathophysiology of von Willebrand disease: Therapeutic implications. Thromb Res. 2011;128(Suppl 1):S3-7.
10. Schnepfenreich 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.
11. 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(10):2103-2114.
12. 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(3):512-520.
13. Ng C, Motto DG, Di Paola J. Diagnostic approach to von Willebrand disease. Blood. 2015;125(13):2029-2037.
14. Ahmad F, Budde U, Jan R, et al. Phenotypic and molecular characterisation of type 3 von Willebrand disease in a cohort of Indian patients. Thromb Haemost. 2013;109(4):652-660.
15. Álvarez‐Laderas I, Nuñez R, Jiménez‐Bárcenas R, et al. The spectrum of mutations in Southern Spanish patients with von Willebrand disease. Haemophilia. 2015;21(3):e240-e242.
16. Yadegari H, Driesen J, Pavlova A, Biswas A, Hertfelder H-J, 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. Thromb Haemost. 2012;108(4):662-671.
17. James PD, Lillicrap D, Mannucci PM. Alloantibodies in von Willebrand disease. Blood. 2013;122(5):636-640.
18. 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. J Thromb Haemost. 2005;3(12):2619-2626.
19. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-249.
20. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS ONE. 2012;7(10):e46688.
21. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep‐sequencing age. Nat Methods. 2014;11(4):361-362.
22. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I‐TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12(1):7-8.
23. Zhou Y-F, Springer TA. Highly reinforced structure of a C‐terminal dimerization domain in von Willebrand factor. Blood. 2014;123(12):1785-1793.
24. Brondijk T, Bihan D, Farndale RW, Huizinga EG. Implications for collagen I chain registry from the structure of the collagen von Willebrand factor A3 domain complex. Proc Natl Acad Sci USA. 2012;109(14):5253-5258.
25. Krieger E, Vriend G. YASARA View - molecular graphics for all devices - from smartphones to workstations. Bioinformatics (Oxford, England). 2014;30(20):2981-2982.
26. Elayaperumal S, Fouzia NA, Biswas A, et al. Type-3 von Willebrand disease in India-Clinical spectrum and molecular profile. Haemophilia. 2018;24(6):930-940.
27. 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.
28. Barocian L, Cozzi G, Canciani MT, et al. Molecular defects in type 3 von Willebrand disease: Updated results from 40 multiethnic patients. Blood Cells Mol Dis. 2003;30(3):264-270.
29. Sutherland MS, Keeney S, Bolton-Maggs P, Hay C, 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. 2009;15(5):1048-1057.
30. Tjernberg P, Castaman G, Vos HL, Bertina RM, Eikenboom J. Homozygous C2362F von Willebrand factor induces intracellular retention of mutant von Willebrand factor resulting in autosomal recessive severe von Willebrand disease. Br J Haematol. 2006;133(4):409-418.

31. 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.

32. Yin J, Ma Z, Su J, et al. Mutations in the D1 domain of von Willebrand factor impair their propeptide-dependent multimerization, intracellular trafficking and secretion. J Hematol Oncol. 2015;8:73.

33. Gupta PK, Saxena R, Adamtziki E, et al. Genetic defects in von Willebrand disease type 3 in Indian and Greek patients. Blood Cells Mol Dis. 2008;41(2):219-222.