

Original Article

Identification and structural characterization of *F8* gene variants in Indonesian patients with nonsevere hemophilia A: Evidence from targeted next generation sequencing and in silico analyses

Niken SN. Handayani^{1*}, Suprianto Suprianto^{2,3}, Ni GAGC. Kirana¹, Indra Lesmana¹, Usi Sukorini⁴, Nastiti Wijayanti¹ and Rarastoeti Pratiwi¹

¹Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, Indonesia; ²Biomedical Sciences and Biotechnology Program, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; ³Center of Excellence for Medical Genomics, King Chulalongkorn Memorial Hospital, Bangkok, Thailand; ⁴Department of Clinical Pathology and Laboratory Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia

*Corresponding author: niken_satuti@ugm.ac.id

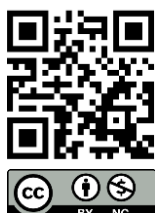
Abstract

Hemophilia A is a rare genetic disorder caused by deficiency of coagulation factor VIII resulting from pathogenic variants in the *F8* gene. The aim of this study was to identify potential pathogenic variants using targeted next generation sequencing (NGS), followed by in silico structural and pathogenicity analyses. Genomic DNA extracted from saliva samples of 11 patients with nonsevere hemophilia A was analyzed using targeted sequencing of *F8* exons 23–26, while targeted NGS was performed on two selected patients for broader variant detection. Identified variants were further evaluated using computational pathogenicity prediction and three-dimensional protein structural modeling. No variants were detected within exons 23–26. Nevertheless, targeted NGS identified a missense variant, c.2095A>G (p.Met699Val), in exon 13 in patient HM1 and a splice-site variant, c.1271+1G>A, in intron 8 in patient HM7. Variant assessment predicted that the variants were classified as likely pathogenic and pathogenic, respectively. Structural comparison between wild-type and mutant factor VIII proteins showed no major differences in overall conformation, secondary structure, or ProMotif features. Nonetheless, physicochemical and atomic interaction analyses indicated local structural perturbations that may affect protein stability and function. These findings suggest that the identified *F8* variants, particularly p.Met699Val, may contribute to factor VIII dysfunction despite preservation of the global protein structure. However, further validation using molecular dynamics simulations and functional experimental studies is required to clarify its structural impact and confirm pathogenicity.

Keywords: Hemophilia A, *F8* gene, variants, factor VIII, in silico

Introduction

Hemophilia A (HA) is a rare inherited bleeding disorder caused by a deficiency of coagulation factor VIII (FVIII) [1]. Based on plasma FVIII levels, HA is classified as mild (>5%–<40%), moderate (1–5%), or severe (<1%) [2]. The variability in disease severity is largely associated with the type and location of variants in the *F8* gene, which play a critical role in determining clinical manifestations [3,4]. The *F8* gene is located on chromosome Xq28, spans approximately 186 kb,



and consists of 26 exons and 25 introns encoding a 2351-amino acid FVIII protein with the domain structure A1-A2-B-A3-C1-C2 [3,5]. Notably, exons 23–26 encode the C1 and C2 domains, which are essential for phospholipid membrane binding and structural stability of FVIII. Several variants in this region have been reported, including missense variants in exon 23 (c.6545G>T-p.R2182L; c.6577T>C-p.C2193R), a frameshift variant in exon 24 (c.7013del-p.L2338Rfs*42), and a deletion in exon 26 (c.6672_6674del-p.S2225del) [6]. Furthermore, variants in exons 23–25 have been shown to disrupt disulfide bond formation, leading to protein destabilization and misfolding [7], thereby potentially altering the structure and function of the C1 and C2 domains.

Globally, the variant spectrum of *F8* is dominated by intron 22 and intron 1 inversions, as well as other variants such as deletions, insertions, splicing, missense, and nonsense variants [8–10]. However, based on major global databases such as ClinVar, the Human Gene Mutation Database (HGMD), the Leiden Open Variation Database (LOVD), and the European Association for Hemophilia and Allied Disorders (EAHAD) database, there have been no reported cases of *F8* variants in patients with HA from Indonesia. This lack of representation highlights a significant gap in global genomic data and underscores the need for population-specific studies. Investigating nonsevere HA is particularly relevant because missense variants are a common molecular cause of mild and moderate HA [3,4] and may contribute to variable clinical phenotypes through their effects on protein structure and function. However, the genetic variation within exons 23–26, a functionally important region of *F8*, has not been specifically characterized in nonsevere HA cases from the Indonesian population. Therefore, the aim of this study was to characterize *F8* variants in patients with nonsevere HA from Yogyakarta, Indonesia, using targeted sequencing of exons 23–26, which encode the C1 and C2 domains of FVIII, complemented by targeted next generation sequencing (NGS) in two selected samples to identify variants beyond these regions. The potential pathogenicity and structural consequences of the identified variants were further evaluated using *in silico* analyses. By providing population-specific genetic evidence, this study may expand the representation of Indonesian patients in global variant databases and inform future molecular diagnosis and genotype–phenotype investigations of nonsevere HA.

Methods

Study design and setting

An exploratory cross-sectional molecular genetic study was conducted among patients with nonsevere HA registered with the Indonesian Hemophilia Society in the Special Region of Yogyakarta, Indonesia. Saliva collection, DNA extraction, targeted sequencing of *F8* exons 23–26, and targeted NGS were conducted at Universitas Gadjah Mada, Yogyakarta, Indonesia. Targeted sequencing was performed for all participants, whereas targeted NGS was performed on two selected samples with sufficient DNA quantity and quality. The identified variants were subsequently validated using Sanger sequencing and evaluated through *in silico* pathogenicity prediction and structural modeling.

Study participants

This study included 11 male patients, designated as HM1-HM11, with nonsevere HA from unrelated families who were registered with the Indonesian Hemophilia Society. Participants were aged 7–22 years and clinically diagnosed with mild to moderate HA based on FVIII activity levels and clinical evaluation. Inclusion criteria included patients diagnosed with mild to moderate HA, registered with the Indonesian Hemophilia Society in Yogyakarta, and willing to participate in the study. Patients with other hematological disorders or incomplete clinical information were excluded.

Saliva collection and DNA extraction

Saliva was used as an alternative DNA source to address ethical considerations in the pediatric population and minimize invasive procedures instead of peripheral blood. This non-invasive approach is more acceptable and safer for pediatric patients. Saliva samples were collected using a Wuxi NEST Dry Saliva Collection Kit (Wuxi NEST Biotechnology Co., Ltd., Jiangsu, China)

according to the manufacturer's instructions. Participants were instructed to avoid eating, drinking, smoking, or chewing gum for at least 30 minutes before sample collection. Prior to collection, participants rinsed their mouths with clean running water to minimize contamination. Approximately 2 mL of saliva was collected into sterile collection tubes containing stabilization buffer and homogenized by gentle inversion. Samples were transported to the laboratory and stored at 4°C until further processing.

For genomic DNA extraction, saliva samples were vortexed thoroughly, and 200 µL of each sample was transferred into a sterile microcentrifuge tube. DNA extraction was performed using the FavorPrep™ 96-Well Genomic DNA Extraction Kit (Favorgen Biotech Corp., Ping-Tung, Taiwan) following the manufacturer's protocol. Genomic DNA concentration and purity were assessed using a spectrophotometer (MaestroNano Micro-Volume Spectrophotometer, MaestroGen Inc., Taiwan) by measuring the A260/A280 ratio. Extracted DNA was stored at -20°C.

DNA amplification

Amplification of *F8* exons 23–26 was performed using polymerase chain reaction (PCR) with a thermal cycler (Bio-Rad, Hercules, CA, USA). Each PCR reaction was performed in a final volume of 25 µL containing 1× PowerPol 2× PCR mix with dye, 0.2 µM of each forward and reverse primer, less than 300 ng of genomic DNA template, and nuclease-free water. Primer sequences, adopted from a previous study [3], used in this study are presented in **Table 1**. PCR amplification was performed with an initial denaturation step at 95°C for 5 minutes, followed by 35 cycles of denaturation at 95°C for 30 seconds, annealing at 59°C for 30 seconds, and extension at 72°C for 30 seconds, followed by a final extension at 72°C for 5 minutes. The amplified products were subsequently maintained at 4–12°C until further analysis.

Table 1. Primers used for amplification of *F8* exons 23–26

Primer	Sequence	Product size (bp)
Exon 23		
Forward	5'-CTTCACTTGCCCCAGACCTAAT-3'	512
Reverse	5'-CCCAGGACTATGCTGGTTTATAGC-3'	
Exon 24		
Forward	5'-TGCAAAAGTTAAAACCTGAGAAATG-3'	422
Reverse	5'-GTCTGCCCATACCAAACTTCC-3'	
Exon 25		
Forward	5'-AGAGTGAGAAGTGCTGTGGTATGG-3'	446
Reverse	5'-AAAGTCACTGTGTTCTCTCAGAATG-3'	
Exon 26		
Forward	5'-TCCCAGATGCGTAGGACAGAGT-3'	398
Reverse	5'-AGCACAAAGGTAGAAGGCAAGC-3'	

DNA sequencing

PCR products of *F8* exons 23–26 were purified and subjected to Sanger sequencing at the Integrated Research and Testing Laboratory (LPPT), Universitas Gadjah Mada, Yogyakarta, Indonesia. Sequencing reactions were prepared using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. Sequencing products were subsequently purified using the BigDye X-Terminator™ Purification Kit (Applied Biosystems, Foster City, CA, USA) and analyzed using the ABI 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Chromatograms from Sanger sequencing were visualized using FinchTV version 1.4.0 (Geospiza, Seattle, WA, USA) to identify variants based on the reference sequence NM_000132.4.

Long-read sequencing and variant confirmation

Targeted NGS was performed using the PromethION 24 platform (Oxford Nanopore Technologies, Oxford, UK) at the Faculty of Biology, Universitas Gadjah Mada, Indonesia. Samples were selected with the highest DNA concentration to support targeted NGS. High-molecular-weight genomic DNA was prepared using the Rapid Barcoding Kit V14 (Oxford Nanopore Technologies, Oxford, UK) following the manufacturer's protocol, including transposase-based fragmentation, adapter ligation, and barcode incorporation.

Barcoded libraries were pooled and loaded onto a PromethION flow cell for sequencing. Raw electrical signal data were basecalled using Oxford Nanopore high-accuracy basecalling models and demultiplexed to generate FASTQ files. Data processing and quality assessment were performed using EPI2ME software (Oxford Nanopore Technologies, Oxford, UK). Reads were aligned to the human reference genome (GRCh38) using the long-read aligner minimap2, and variant calling was performed using Clair3, as implemented in the EPI2ME human-variation pipeline v2.2.2, to identify single-nucleotide variants and small insertions/deletions. Variants located within the *F8* gene region on chromosome X (chrX: 154,835,792–155,022,723; GRCh38) were prioritized for further analysis.

Variants detected by targeted NGS were validated using PCR amplification followed by Sanger sequencing. The Sanger sequencing data were analyzed using the Benchling platform version 2025 (Benchling Inc., San Francisco, CA, USA). The primers used were F: 5'-TGGGGCCCAACTGATTATTA-3' and R: 5'-TCAGAGGAACTAGATCCCTGTA-3', and F: 5'-TGAAATGGATGTGGTCAGGTT-3' and R: 5'-AAAGGTCCCAAGATTCCTGA-3'. The primers were designed using Primer3Plus (<https://www.primer3plus.com>) based on the reference sequence NM_000132.4.

In silico variant assessments

Variant characterization and pathogenicity interpretation were performed using the Franklin platform by Genoox (Qiagen Digital Insights). In this study, pathogenicity interpretation involved assessing the potential functional impact of the identified variants using multiple computational predictions, including missense effect, deleteriousness, and potential effects on pre-mRNA splicing. These assessments incorporated predictions from AlphaMissense, VARIETY, DANN, MetaLR, SpliceAI, dbSNV, and BayesDel, which were integrated by the Franklin platform to support the overall interpretation of variant pathogenicity [11,12]. Additional assessments were performed using MutationTaster to further evaluate the predicted functional effects of identified variants [13] and the gnomAD Browser v4.1.1 to assess their population allele frequencies and rarity [14].

Protein structure and physicochemical analysis

The three-dimensional (3D) structure of the FVIII mutant protein (p.Met699Val) was constructed using the Phyre2 protein homology/analogy recognition engine version 2.0 (<https://www.sbg.bio.ic.ac.uk>; Job ID: caed8cof252d60e7) based on the amino acid reference sequence NP_000123.1, following the protocol described previously [15,16]. Secondary structure (2D) elements and ProMotif features were evaluated using PDBsum (<https://www.ebi.ac.uk>). The generated protein models were subsequently visualized using Discovery Studio Visualizer v24.1.0.23298 (Dassault Systèmes BIOVIA, San Diego, CA, USA).

Physicochemical properties of the wild-type and mutant FVIII proteins, including molecular weight, theoretical isoelectric point, instability index, aliphatic index, and grand average hydropathicity (GRAVY), were calculated using ProtParam (ExPASy Bioinformatics Resource Portal; <https://web.expasy.org/protparam/>).

Prediction of protein stability and structural impact

Protein stability changes induced by amino acid substitutions were analyzed using DynaMut2 version 2.0 (<https://biosig.lab.uq.edu.au>), which predicts variant-induced changes in stability and flexibility. The structural and functional effects of the p.Met699Val variant were further analyzed using the HOPE server version 2025 (<https://www3.cmbi.umcn.nl/hope/>) following the protocol described previously [17].

The potential damaging effects of amino acid substitutions were further assessed using PolyPhen-2 version 2.2.3 (<https://genetics.bwh.harvard.edu>). Atomic interactions and residue-level structural changes between wild-type and mutant FVIII proteins were analyzed using UCSF Chimera version 1.19 (Resource for Biocomputing, University of California, San Francisco, CA, USA).

Results

Detected variants by targeted NGS

Variability in DNA yield and quality was observed across samples, and not all saliva samples met the required concentration and purity thresholds for targeted NGS. Consequently, two samples were selected for further genomic analysis. Considering the quantity and quality of the DNA samples, HM1 (28.4 ng/ μ L; A260/280=1.98) and HM7 (4.2 ng/ μ L; A260/280=1.28) were selected for targeted NGS to support reliable variant detection and minimize the risk of suboptimal sequencing results. The analysis detected a missense variant c.2095A>G (chrX:154947716 T>C) in patient HM1 and a splice-donor variant c.1271+1G>A (chrX:154966425 C>T) in patient HM7, the latter of which is reported in ClinVar (rs1603435260) as pathogenic.

The identified variants were subsequently validated using Sanger sequencing in the respective patients. In addition, nine other patients were screened for these variants. The results confirmed that the c.2095A>G variant was detected only in HM1 and the c.1271+1G>A variant only in HM7 (**Figure 1**), with no evidence of these variants in the remaining samples.

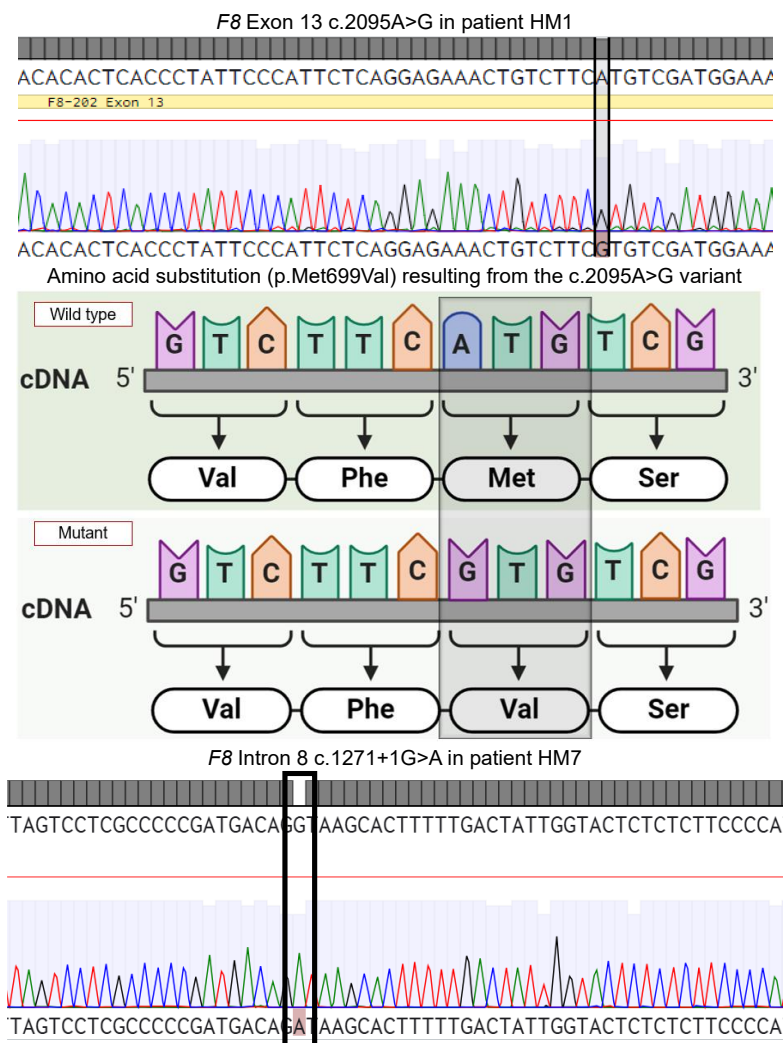


Figure 1. *F8* gene variants in exon 13 and intron 8 detected in nonsevere hemophilia A patients.

Variant assessment

The c.2095A>G missense variant was predicted by AlphaMissense to have an uncertain impact (score=0.444). VARIETY classified this variant as deleterious (score=0.71), supported by DANN (score=0.97) and MetaLR (score=0.98). For the c.1271+1G>A splice-site variant, SpliceAI predicted a splicing impact with a score of 0.99. Consistently, dbSNV AdaBoost classified this variant as deleterious (score=1), and BayesDel (no AF) classified it as deleterious (strong) with a score of 0.83, exceeding the ClinGen threshold (>0.5). Identified *F8* variants were evaluated using the ACMG/AMP sequence-variant interpretation framework through the Franklin by

Genoox platform. Evidence categories considered included population frequency (PM2), variant location and functional relevance (PM1), computational predictions (PP3), effects on protein function or splicing (PVS1), and previously reported variation at the same amino acid position or related clinical evidence (PM3, PM5, PP2, and PP5), where applicable. The combined evidence was used to assign the final ACMG/AMP classification. *F8* c.2095A>G (p.Met699Val) was classified as likely pathogenic, whereas *F8* c.1271+1G>A was classified as pathogenic, respectively (**Figure 2**). MutationTaster predicted both variants to be deleterious, indicating their potential pathogenic effects on protein function, as presented in **Table 2**.

Table 2. Description of detected *F8* gene variants in patients with nonsevere HA

Variant	Prediction	DNA changes	AA changes	Splice sites	Features at a glance
Exon 13					
chrX:154947716 T>C	Deleterious	c.2095A>G g.79225A>G	p.Met699Val	No abrogation of potential splice sites predicted by MaxEntScan	Amino acid sequence changed, protein features (might be) affected
Intron 8					
chrX:154966425 C>T	Deleterious	c.1271+1G>A g.60516G>A	N/A	Donor lost 60515 CCGATGACAG ataagcactt	ClinVar ID 811704 (pathogenic) Splice site lost



Figure 2. In silico pathogenicity classification of *F8* gene variants identified in patients with nonsevere hemophilia A: (A) chrX:154947716T>C (c.2095A>G) and (B) chrX:154966425C>T (c.1271+1G>A).

In the integrated Ensembl Variant Effect Predictor (VEP) annotation within gnomAD, the *F8* exon 13 c.2095A>G variant identified in patient HM1 was absent from the database. The *F8* intron 8 c.1271+1G>A variant (SNV: X-154966425 C>T, GRCh38) was reported with a global allele frequency of 1.65×10^{-6} and was identified only in African/African American and non-Finnish European populations, with a single allele reported in each population. The same variant was also identified in a Javanese patient with nonsevere HA from Indonesia in this study. VEP annotation in gnomAD classified this splice-site alteration as a high-confidence predicted loss-of-function (pLoF) variant.

Structural and physicochemical analysis of the p.Met699Val FVIII variant

The 3D structural models showed no significant differences in the overall surface conformation between wild-type and mutant FVIII proteins (**Figure 3**). In addition, secondary structure analysis using PDBsum shows no differences in structural elements, such as β -sheets and loops. ProMotif analysis was also consistent, with both wild-type and mutant proteins displaying the same composition: three sheets, one β -hairpin, two β -bulges, eight strands, six β -turns, and one

γ-turn. These findings indicate a similar global structural architecture. The physicochemical properties of the wild-type and mutant FVIII proteins are presented in **Table 3**.

Table 3. Comparison of the physicochemical properties of wild-type and mutant FVIII proteins

Type	Molecular weight (kD)	Formula	Total number of atoms	Instability index	Aliphatic index	Grand average hydropathicity (GRAVY)
Wild-type	8034.18	C ₃₇₇ H ₅₄₃ N ₈₁ O ₁₀₆ S ₄	1111	31.37	83.43	0.359
Mutant	8002.12	C ₃₇₇ H ₅₄₃ N ₈₁ O ₁₀₆ S ₃	1110	25.09	87.57	0.391

A

KEY = DOMAINS: Signal Peptide- A1 - a1 - A2 - a2 - B - a3 - A3 - C1 - C2

MQIELSTCF	LCLLRFCFA	TRRYLGA	LSWDYMQSD	GELPVDARFP	50	IVDDTSTQWS	KNMKHLTPST	LTQIDYNEKE	KGAITQSPLS	DCLTRSHSIP	1400
PRVPSFFPN	TSVYKTKLF	VEFTDHLFNI	AKPRPPWML	LGPTIQAEVY	100	QANRSPLEPIA	KVSSFPFSIRP	IYLRVLFQD	NSSHLPAASY	RKKGSGVQFP	1450
DTVVITLQNM	ASHFVSLHAV	GVSYNKASEG	AEYDDQTSQR	EKEDDKVFPY	150	SHFLQGAKKN	NLSLAITLVE	MTGDQREVGS	LGTSATNSVT	YKKVENTVLP	1500
GSHTYVQVVL	KENGPMASDP	LCLTYSLSH	VDLVKDLNSG	LIGALLVCRE	200	KPDLPTKSGK	VELLPKVHIY	QKDLFPETST	NGSPGHLDLV	EGSLQGTETG	1550
GSIAKEKTQT	LHKFILLFAV	FDEGKSWHSE	TKNSLMQDRD	AASARAWPKM	250	AIKWNEANRP	GKVPFLRVAT	ESSAKTPSKL	LDPLAWNDHY	GTQIPKEWK	1600
HTVNGVNMRS	LPFLIGCHRK	SVYHNVIGMG	TPEVHSIFL	EGHTFLVRNH	300	SQEKSPKTA	FKKKTILSL	NACENSHAIA	AINEGQNKPE	IEVTWAKQGR	1650
RQASLEISPI	TFIATQITLM	DLGQFLFCH	ISSHQHDMGE	AYKVDSCPE	350	TERLCSQNP	VLKRHRQREIT	RTTLQSDQEE	IDYDDTISVE	MKEDFDIYD	1700
EPQLRMKNNE	EAEDYDDDLT	DSEMDVVRFD	DNNSPSTFIQI	RSVAKKHFKT	400	EDENQSPRSF	QKTRHYFIA	AVERLDWYDM	SSSPHVLNRN	AQSGSVQFPK	1750
WVHYIAAEE	DWDYAPLVLA	PDRSYKSYQY	LNNGPQRIGR	KYKVRFMAY	450	KVVFQEPFDG	SFTQFLIRGE	LMHGLGLLP	YIRAEVEDNI	WVTFRNQASR	1800
TDETFFKTREA	IQHESGILGP	LLYGEVGDITL	LIIFKNQASR	PYNIYPHGIT	500	PYSFYSLIS	VEEDQROGAE	PRKNFVKPNE	TKTYFWKQCH	HMATPKQEF	1850
DVRELYSRL	PKGVKHLKDF	PILPGEIFKY	KWTVIVEDGP	TKSDPRLCTR	550	LFDETSKWF	TENMERNCR	PCNIQMEDPT	FKENYRPHAI	NGYIMDLPG	1900
YSSFFVNER	DIAAGLIGPL	LCKVKSVDQ	RGNQIMSDR	NVLPSVFTS	600	LVMADQDRI	WYLLSMGSE	NIHSIFSGH	VFTVRKKEEY	KMALYNLYPG	1950
NRSWYLTENI	QRLPNPAGV	QLEDPEFQAS	NIMHSINGV	PDLQLQSVCL	650	VEETVEMLES	KAGIWRVECL	IGEHLAGMS	TLFLVSNKRC	QTPLMASGH	2000
HEVAYWYLS	IGAQTDFLSV	FFSGYTFKHK	MYVEDTLTLE	PFSGETVFMEN	700	IRDFQITASG	QYQWAPKLA	RLHYSGSINA	WSTKEPFSWI	KVDLLAPMII	2100
MENGLMILG	CHNSDFNRNG	MTALLKVSSC	DKNTGDYED	SYEDISAYLL	750	HGKIQGARG	KFSSSLYSQF	LIMYSLDGKK	WQTYRGNSTG	TLMVFFGNVD	2150
SKNNAIEPR	FSQNSRHSPT	RQKQFNATTI	PENDIEKTEP	WFAHRTMEPK	800	SGGIKHINFI	PPIIARYIRL	HPHTYSIRST	LMNLMGCDL	NSCSMPLGME	2200
IGNVSSDILL	MLLRQSPPTP	GLSLDLQEA	KYETFSDDPS	PGADISNNSL	850	SKAISDAQIT	ASSYFTNMPA	TWSPSKARLH	LQGRSNAMRP	QVNNPKWELQ	2250
SEMTHTSRQL	HHSQDMVFTP	ESGLQLRLNE	KLGTTAATLE	KKLDKFSYST	900	VDFQTKMVT	GVTTFQGVKSL	LTSMYVKEPL	ISSSQDGHQW	TLFFNGKGVK	2300
SNMLISTPIS	DNLAAGTNT	SSGLPSPMPV	HYDSQDLTTL	FGKSSPLTE	950	VFOGNDSET	PVNVSLDPL	LTTRYLRHPQ	SWHVIQALRM	EVLGCEAODL	2350
SGGLSLSEE	NNSDKLLES	LMNSQESSWG	KNVSSSTESR	LFGKGRAPGH	1000						
ALLTKNALF	KVISLTKTN	KTSNNSATNR	KTHIDGFSLL	IENSPPVQNW	1050						
ILESDFEKK	VTPLIHDRML	MDKNATALRL	NHMSNKTTS	KNMEMVQKK	1100						
EGPIPPDAQN	PDMSFFKMLF	LPESARWIQR	THGKNSLNSG	QGPSPQLVLS	1150						
LGPEKSEVGQ	NFLSEKNKVV	VKGGEFTKDV	GLKEMVPSS	RNLFLTNLDN	1200						
LRNNTHNQE	KKIQEEIEKK	ETLIQENVVL	PQIHTVTGK	NFMKNLFLS	1250						
TRQNVESGYD	GAYAPVLQDF	RSNDSTNRT	KKHTAHFSKK	GEENLEGGL	1300						
NQTKQIVEKY	ACTRISPNF	SQNFVTVQRS	KRALQFRLP	LEETELEKRI	1350						

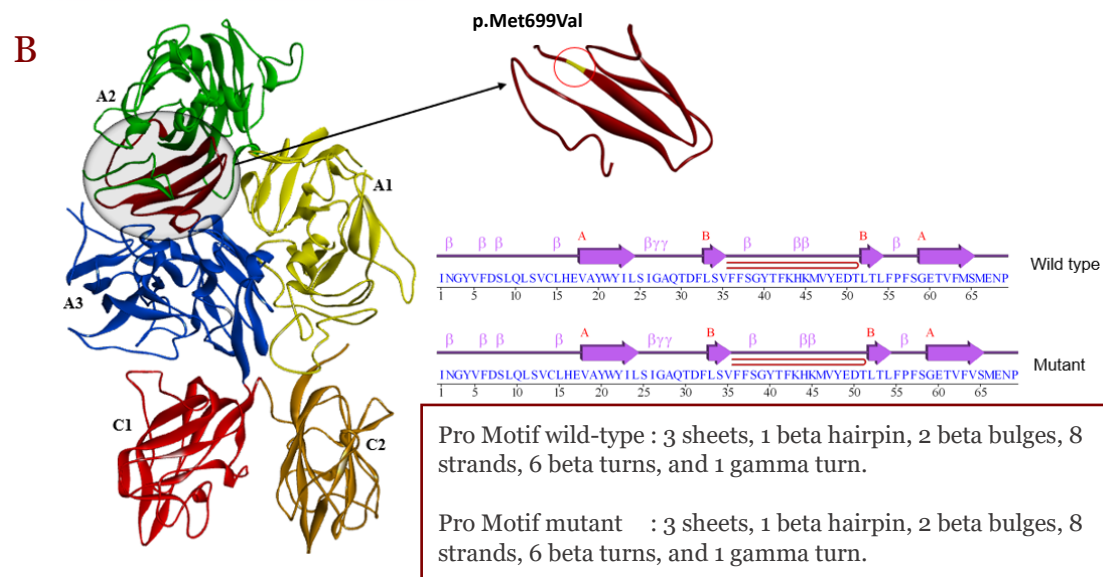


Figure 3. Three-dimensional structural analysis of the mutant FVIII protein carrying the p.Met699Val substitution. (A) Amino acid residues highlighted in red represent previously reported point mutation variants, while the yellow-highlighted region within the A2 domain indicates the target amino acid sequence encoded by exon 13. (B) Structural comparison between the p.Met699Val mutant and wild-type FVIII proteins showed no observable differences in secondary structure organization or ProMotif features.

Structural stability and atomic interaction analysis of the p.Met699Val FVIII variant

Structural analysis showed the local environment of the mutated residue at position 699 in chain A (**Figure 4**). The methionine-to-valine substitution was located within the highlighted region and was associated with changes in local atomic interactions. The interaction network surrounding the residue, including hydrogen bonds and other atomic contacts, showed differences between the wild-type and mutant models. Stability prediction using DynaMut2 yielded a $\Delta\Delta G$ value of -0.43 kcal/mol, indicating a predicted destabilizing effect.

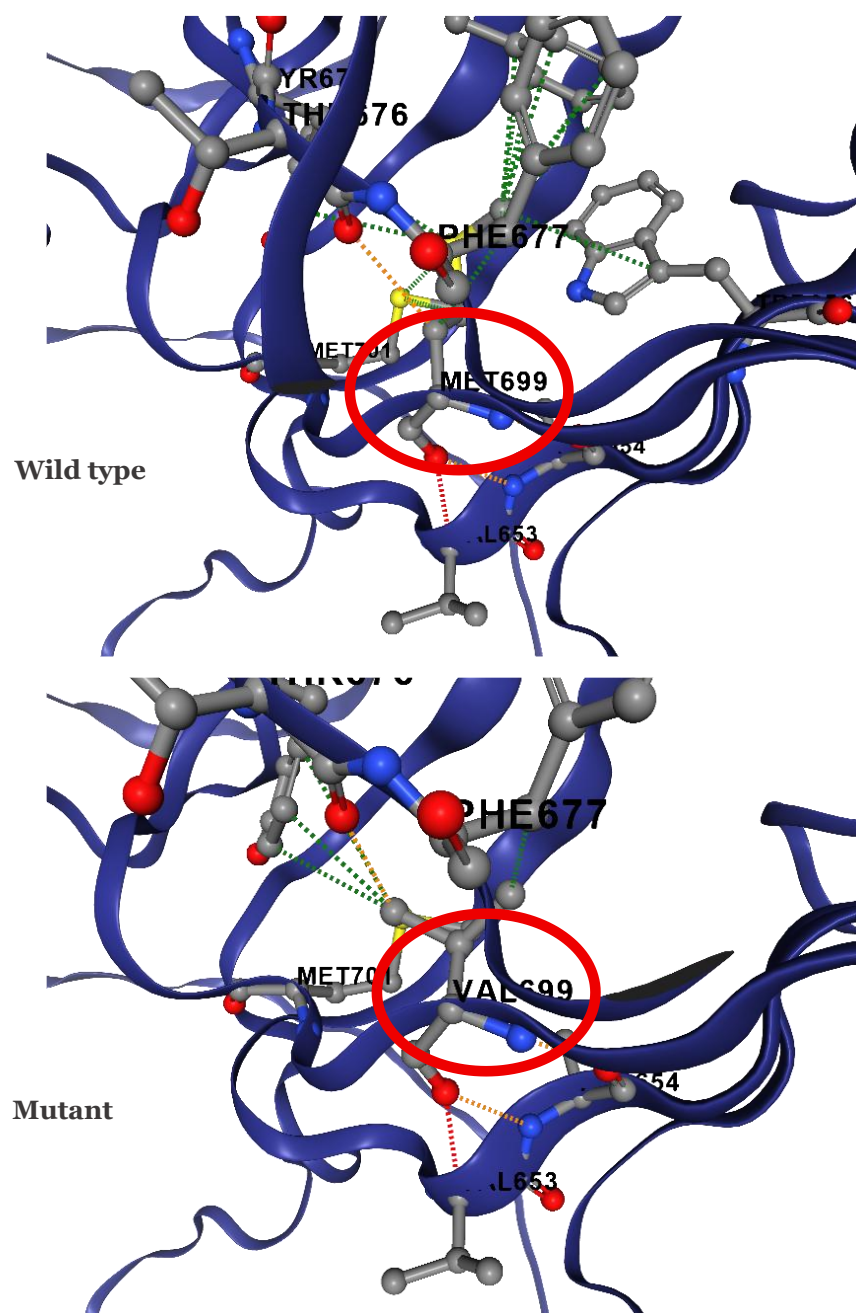


Figure 4. Structural and atomic interaction analysis of FVIII carrying the p.Met699Val substitution, showing altered atomic interactions at residue 699, with red circles highlighting the mutated residue and dashed lines indicating predicted atomic interactions.

Structural analysis of the p.Met699Val variant showed the local conformation of residue 699 in the wild-type (green) and mutant (red) structures across multiple viewing angles (**Figure 5**). The methionine-to-valine substitution resulted in differences in side-chain orientation and spatial positioning within the highlighted region. Structural overlay demonstrated a slightly altered conformation of the mutant residue within the same structural region, while no apparent changes in the overall backbone structure were observed.

Structural and sequence-based analyses of the variant and its atomic interactions are presented in **Figure 6**. Secondary structure prediction showed a localized difference in the highlighted region between the wild-type and mutant structures. The 3D structural mapping highlighted the location of the variant within the protein, as indicated by the circled region. Local interaction analysis showed differences in atomic contacts between the wild-type and mutant structures within this region. In addition, the variant was classified as probably damaging, with a prediction score of 1.000.

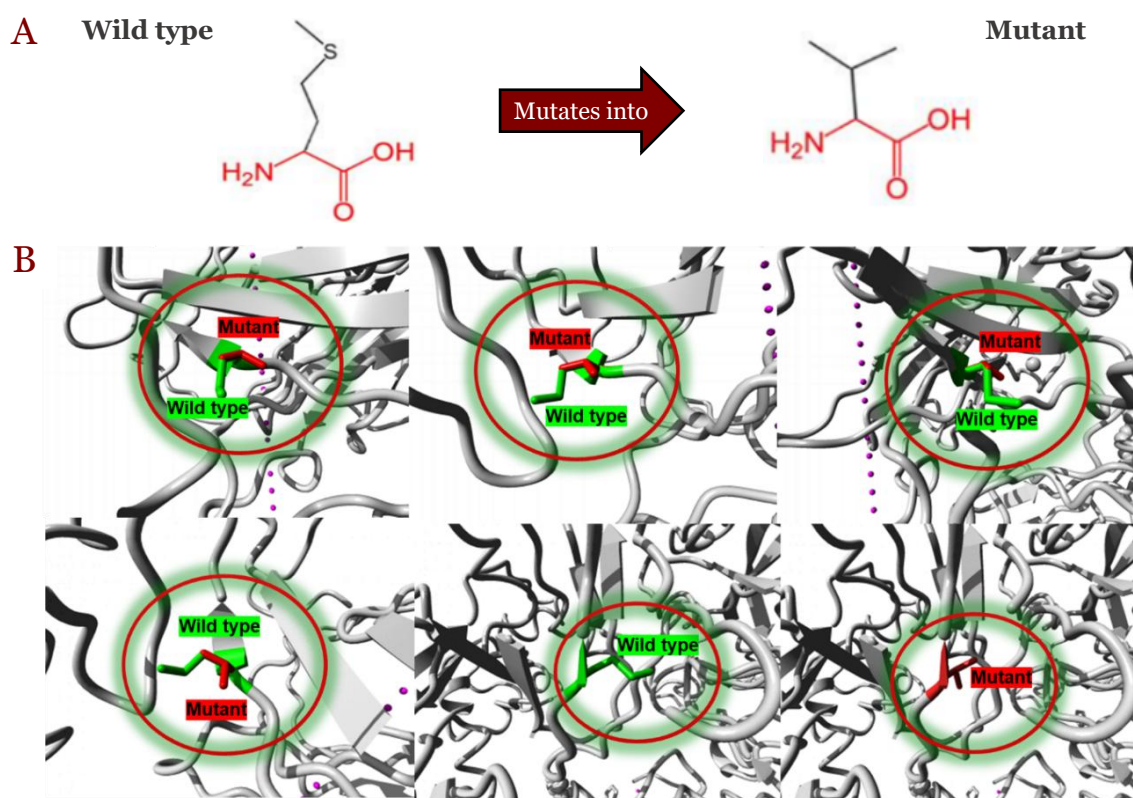


Figure 5. Predicted structural effects of the p.Met699Val variant on FVIII. (A) Comparison of the physicochemical properties between the wild-type and mutant residues, showing that the mutant residue is smaller than the wild-type residue. (B) Close-up visualization of the mutation site, with the wild-type and mutant side chains shown in green and red, respectively.

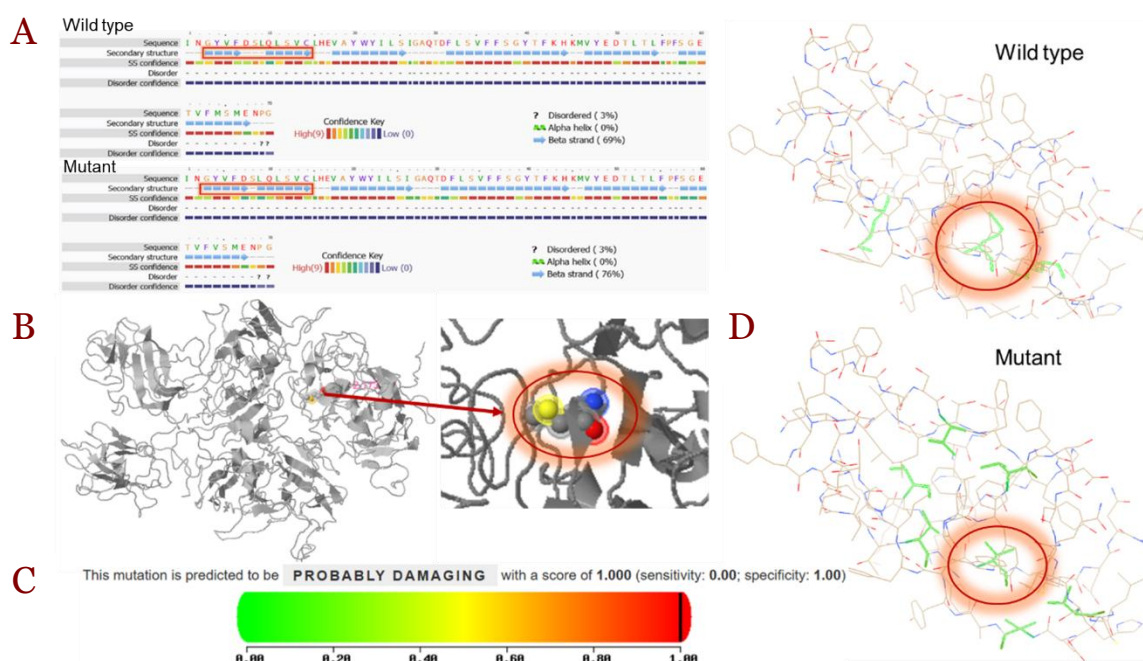


Figure 6. Structural and functional prediction analysis of the FVIII p.Met699Val variant. (A) Two-dimensional structural representation of the mutant FVIII protein (p.Met699Val) compared with the wild-type FVIII. (B) Three-dimensional visualization of the predicted damaging region. (C) PolyPhen-2 prediction indicating a probably damaging effect of the variant on protein function. (D) Comparison of atomic interactions between wild-type and mutant FVIII proteins in the region encoded by exon 13.

Discussion

Exploration of *F8* gene variants in nonsevere HA remains important because some cases cannot be fully explained by classical structural variants, such as intron 22 and intron 1 inversions. This has directed attention toward other potentially relevant regions, particularly exons 23–26, which encode the C1 and C2 domains of FVIII and represent mutationally active regions of the *F8* gene [5]. Although previous studies have reported pathogenic alterations in these exons, including duplications [4,6,17], deletions [18], nonsense [4,19–21], and missense [22], no variants were identified in the present study. These findings suggest that the molecular basis of HA in these patients may involve variants outside the analyzed C1 and C2 domain–coding regions, including other coding exons, intronic or regulatory regions, or structural abnormalities such as intron 22 and intron 1 inversions.

Consistent with this possibility, NGS in the present study identified variants in exon 13 (c.2095A>G) and intron 8 (c.1271+1G>A) of the *F8* gene. The c.2095A>G variant is a missense change predicted by AlphaMissense to have an uncertain impact, with a score of 0.444. This value falls within the intermediate range, indicating that it is neither strongly associated with pathogenic nor benign effects [23]. In contrast, VARIETY predicts c.2095A>G as deleterious (score=0.71), suggesting a high likelihood of functional impact [24]. This interpretation is supported by DANN (0.97) and MetaLR (0.98), both of which predict a deleterious effect, indicating a potential adverse impact on protein function [25,26]. For the c.1271+1G>A variant, splicing prediction tools consistently indicate a strong functional impact. Predictions from SpliceAI suggest a high probability that this variant alters normal RNA splicing, as scores closer to 1 (0.99) indicate a greater likelihood of splice disruption [27]. In parallel, dbcsSNV AdaBoost classifies this variant as deleterious (score=1), representing the highest confidence that the variant negatively affects splicing within consensus regions [28]. Consistently, BayesDel (no AF) further supports a damaging effect, classifying the variant as deleterious (strong) with a score of 0.83, which exceeds the ClinGen threshold (>0.5) for strong pathogenic evidence [29]. Collectively, these in silico predictions strongly suggest that the variant may disrupt normal splicing and contribute to impaired gene function.

Population data indicate that both variants are rare. The c.2095A>G variant identified in patient HM1 is absent from large population databases, including gnomAD, indicating limited evidence regarding its frequency and distribution. However, the same variant has previously been reported in patients with moderate HA from Colombia and Turkey [6,30], suggesting that it may represent a rare recurrent variant rather than a population-specific alteration. The similar clinical presentation among reported cases may indicate a possible genotype–phenotype association, although interpretation remains limited because of the small number of reported cases and potential influences of genetic background. Similarly, the c.1271+1G>A variant is classified as ultra-rare in gnomAD, with an allele frequency of 1.65×10^{-6} . Previously, this variant had only been reported in African/African American and non-Finnish European populations. In the present study, the same variant was identified in a Javanese patient with nonsevere HA from Yogyakarta, Indonesia, expanding the currently reported population spectrum. The variant was identified in a hemizygous state in a male patient because the *F8* gene is located on the X chromosome [31], consistent with the X-linked inheritance pattern of HA. The guanine-to-adenine substitution at this splice donor site is predicted to affect splice recognition and may contribute to abnormal splicing events, such as exon skipping, intron retention, or activation of cryptic splice sites.

The potential effects of these variants can also be considered in relation to FVIII domain organization. The A1, A2, and A3 domains function as essential procofactors and share structural homology with ceruloplasmin [32], whereas the C1 and C2 domains mediate interactions with von Willebrand factor (VWF) and phospholipids that are important for FVIII stability and activity [33–35]. Exon 13 encodes part of the A2 domain, which is involved in activated factor IX (FIXa) binding [36]. Therefore, the p.Met699Val substitution may influence local folding or intermolecular interactions associated with FIXa binding, potentially contributing to reduced cofactor activity without complete loss of protein function.

Physicochemical analysis showed that the wild-type and mutant proteins retained generally similar global properties, indicating no major structural alterations. However, subtle differences

may still influence local residue interactions, packing density, stability, and protein dynamics [37]. Increased instability index values suggest reduced structural robustness, whereas changes in aliphatic index, GRAVY values, and charged residues indicate possible alterations in hydrophobicity, thermal behavior, and electrostatic interactions [38-40]. Together, these findings suggest that local structural perturbations may affect protein function despite preservation of the overall fold.

Structural analyses supported the presence of localized changes associated with the p.Met699Val substitution. Replacement of methionine with the smaller valine residue reduces side-chain volume, potentially creating a small cavity and weakening local packing density. The substitution may also alter side-chain orientation and interatomic interaction networks within FVIII, including hydrogen bonding and van der Waals contacts [40-44]. Methionine contains a sulfur atom that can contribute to local structural interactions, whereas valine lacks this feature [45]. Consequently, replacing methionine with valine may alter local interactions and hydrophobic packing within the A2 domain. These subtle structural changes may influence local folding stability or intermolecular interactions without causing major disruption to the overall protein structure. However, these findings are based primarily on computational predictions and static structural models. Further experimental studies and molecular dynamics simulations are needed to determine whether these predicted changes affect FVIII stability, interactions, or biological function.

This study has some limitations that need to be acknowledged. First, the predominantly pediatric cohort necessitated the use of non-invasive saliva sampling to minimize the burden associated with invasive blood collection. This approach resulted in relatively low DNA yield and purity, although the recovered DNA remained sufficient for downstream molecular analyses, including targeted NGS. Second, variant interpretation was based primarily on *in silico* analyses, and the functional consequences of the identified *F8* variants were not experimentally validated. The study was also limited to patients with nonsevere hemophilia A from Yogyakarta, which may limit the generalizability of the findings to other populations and clinical settings. Functional studies are therefore warranted to determine the effects of the identified variants on *F8* expression, pre-mRNA splicing, or protein function, and to further support their clinical interpretation.

Conclusion

No variants were detected in exons 23–26 of the *F8* gene. However, targeted NGS identified a missense variant, c.2095A>G (p.Met699Val), in exon 13 of patient HM1 and a splice-site variant, c.1271+1G>A, in intron 8 of patient HM7. Although the predicted physicochemical properties of the mutant FVIII protein are largely comparable to the wild-type, *in silico* analyses, including damaging-effect prediction and altered atomic interaction profiles, suggest that p.Met699Val may affect local structural stability, residue interactions, and the functional integrity of the A2 domain involved in FIXa binding. The p.Met699Val substitution appears to induce subtle local conformational and interaction changes rather than major global structural disruption, which may contribute to reduced FVIII cofactor activity and the moderate HA phenotype observed in HM1. Nonetheless, these findings are based on computational models and static structural predictions. Further functional and experimental studies are required to confirm the biological and clinical significance of this variant.

Ethics approval

This study was approved by the Medical and Health Research Ethics Committee, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Indonesia (Approval No. KE/FK/1418/EC/2023 on September 1, 2023). Written informed consent was obtained from all participants or their legal guardians prior to sample collection.

Acknowledgments

The authors thank the patients and members of the Indonesian Hemophilia Society, Yogyakarta Special Region, for their participation, and Dr. Pudjo Hagung Widjajanto and Dr. Agus Wibowo for valuable discussions.

Competing interests

All the authors declare that there are no conflicts of interest.

Funding

This study was funded by the Academic Excellence Enhancement Program (Scheme A) for NSNH (Grant No. 4301/UN1/DITLIT/PT.01.03/2024).

Underlying data

Derived data supporting the findings of this study are available from the corresponding author on request.

Declaration of artificial intelligence use

This study used QuillBot solely for language refinement. No artificial intelligence (AI) tools were used for data analysis or interpretation. All data generation, analysis, and conclusions presented in this manuscript are the sole responsibility of the authors. All content was critically reviewed by the authors to ensure the integrity and reliability of the results, and all final decisions and interpretations were made independently by the authors.

How to cite

Handayani NSN, Suprianto S, Kirana NGAGC, *et al.* Identification and structural characterization of *F8* gene variants in Indonesian patients with nonsevere hemophilia A: Evidence from targeted next generation sequencing and in silico analyses. Narra X 2026; 4 (3): e297 - <http://doi.org/10.52225/narrax.v4i3.297>.

References

1. Lim MY, Lee K, Johnsen JM, *et al.* F8 Variants and inhibitor development in a multiethnic cohort of nonsevere haemophilia A. Haemophilia 2025;31(6):1217-1225.
2. Blanchette VS, Key NS, Ljung LR, *et al.* Definitions in hemophilia: Communication from the SSC of the ISTH. J Thromb Haemost 2014;12(11):1935-1939.
3. Wang J, Gu J, Chen H, *et al.* A novel deletion mutation of the F8 gene for hemophilia A. Diagnostics 2022;12(2876).
4. Johnsen JM, Fletcher SN, Huston H, *et al.* Novel approach to genetic analysis and results in 3000 hemophilia patients enrolled in the My Life, Our Future initiative. Blood Adv 2017;1(13):824-834.
5. Gunasekera D, Vir P, Karim AF, *et al.* Hemophilia A subjects with an intron-22 gene inversion mutation show CD4+ T-effector responses to multiple epitopes in FVIII. Front Immunol 2023;14:1128641.
6. Atik T, Işık E, Onay H, *et al.* Factor 8 gene mutation spectrum of 270 patients with hemophilia A: Identification of 36 novel mutations. Turk J Hematol 2020;37(3):145-153.
7. Zimmermann MA, Oldenburg J, Müller CR, *et al.* Characterization of duplication breakpoints in the factor VIII gene. J Thromb Haemost 2010;8(12):2696-2704.
8. Abdulqader AMR, Mohammed AI, Rachid S, *et al.* Identification of the intron 22 and intron 1 inversions of the factor VIII gene in Iraqi Kurdish patients with hemophilia A. Clin Appl Thromb 2020;26:1-7.
9. Pshenichnikova O, Salomashkina V, Poznyakova J, *et al.* Spectrum of causative mutations in patients with hemophilia A in Russia. Genes 2023;14(2):260.
10. Santacroce R, Acquila M, Belvini D, *et al.* Identification of 217 unreported mutations in the F8 gene in a group of 1,410 unselected Italian patients with hemophilia A. J Hum Genet 2008;53(3):275-284.
11. Costa M, García S A, León A, *et al.* The promises and pitfalls of automated variant interpretation: A comprehensive review. Brief Bioinform 2025;26(5):bbaf545.
12. Kaul D, Le H, Kwan E, *et al.* The utility of artificial intelligence for DNA variant curation in the Mendelian genetic disorder familial hypercholesterolaemia. Pathology 2026;58(5):571-578.
13. Steinhaus R, Proft S, Schuelke M, *et al.* MutationTaster2021. Nucleic Acids Res 2021;49(W1):W446-W451.
14. Gudmundsson S, Singer-Berk M, Watts NA, *et al.* Variant interpretation using population databases: Lessons from gnomAD. Hum Mutat 2022;43(8):1012-1030.

15. Kelley LA, Mezulis S, Yates CM, *et al.* The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015;10(6):845-858.
16. Powell HR, Islam SA, David A, *et al.* Phyre2.2: A community resource for template-based protein structure prediction. *J Mol Biol* 2025;437(15):1-17.
17. Venselaar H, te Beek TAH, Kuipers RKP, *et al.* Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics* 2010;11:548.
18. Youssoufian H, Antonarakis SE, Aronis S, *et al.* Characterization of five partial deletions of the factor VIII gene. *Proc Natl Acad Sci U S A* 1987;84(11):3772-3776.
19. Astermark J, Oldenburg J, Escobar M, *et al.* The Malmö International Brother Study (MIBS). Genetic defects and inhibitor development in siblings with severe hemophilia A. *Haematologica* 2005;90(7):924-931.
20. Ravanbod S, Rassoulzadegan M, Rastegar-Lari G, *et al.* Identification of 123 previously unreported mutations in the F8 gene of Iranian patients with Haemophilia A. *Haemophilia* 2012;18(3):340-346.
21. Citron M, Godmilow L, Ganguly T, *et al.* High throughput mutation screening of the factor VIII gene (F8C) in hemophilia A: 37 Novel mutations and genotype-phenotype correlation. *Hum Mutat* 2002;20(4):267-274.
22. Hwang SH, Kim MJ, Lim JA, *et al.* Profiling of factor VIII mutations in Korean haemophilia A. *Haemophilia* 2009;15(6):1311-1317.
23. Cheng J, Novati G, Pan J, *et al.* Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* 2023;381(6664):eadg7492.
24. Wu Y, Li R, Sun S, *et al.* Improved pathogenicity prediction for rare human missense variants. *Am J Hum Genet* 2021;108(10):1891-1906.
25. Quang D, Chen Y, Xie X. DANN: A deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* 2015;31(5):761-763.
26. Dong C, Wei P, Jian X, *et al.* Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* 2015;24(8):2125-2137.
27. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, *et al.* Predicting splicing from primary sequence with deep learning. *Cell* 2019;176(3):535-548.
28. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res* 2014;42(22):13534-13544.
29. Feng BJ. PERCH: A Unified Framework for disease gene prioritization. *Hum Mutat* 2017;38(3):243-251.
30. Yunis LK, Linares A, Cabrera E, *et al.* Systematic molecular analysis of hemophilia A patients from Colombia. *Genet Mol Biol* 2018;41(4):750-757.
31. Morris O, Morris M, Jobe S, *et al.* Genomic landscape of chromosome X factor VIII: From hemophilia A in males to risk variants in females. *Genes* 2024;15(12):1-18.
32. Pemberton S, Lindley P, Zaitsev V, *et al.* A molecular model for the triplicated A domains of human factor VIII based on the crystal structure of human ceruloplasmin. *Blood* 1997;89(7):2413-2421.
33. Print EHT, Simmons AM, Spencer HJ, *et al.* Analyzing 6211 unique variants in the upgraded interactive FVIII web database reveals novel insights into hemophilia A. *Blood Vessels Thromb Hemost* 2025;2(3):100053.
34. Gangadharan B, Ing M, Delignat S, *et al.* The C1 and C2 domains of blood coagulation factor VIII mediate its endocytosis by dendritic cells. *Haematologica* 2017;102(2):271-281.
35. Stoilova-McPhie S, Lynch GC, Ludtke S, *et al.* Domain organization of membrane-bound factor VIII. *Biopolymers* 2013;99(7):448-459.
36. Jiang S, Li F, Li L, *et al.* The construction of a molecular model for the ternary protein complex of intrinsic coagulation pathway factors provides novel insights for the pathogenesis of cross-reactive material positive coagulation factor mutations. *Int J Mol Sci* 2025;26(11):1-19.
37. Yang K, Chen Y, Wei Y, *et al.* MicroEnvPPI: Microenvironment-aware optimization enables generalizable protein-protein interaction prediction. *J Chem Inf Model* 2025;65(21):11860-11877.
38. Midlovets K, Volkova N, Peka M. Bioinformatic analysis of WNT family proteins. *Bioinform Biol Insights* 2025;19:11779322251353347.
39. Rye TK, Martinovic G, Eie LV, *et al.* Electromembrane extraction of peptides using deep eutectic solvents as liquid membrane. *Anal Chim Acta* 2021;1175:338717.
40. Zhou HX, Pang X. Electrostatic interactions in protein structure, folding, binding, and condensation. *Chem Rev* 2018;118(4):1691-1741.

41. Wakabayashi H, Fay PJ. Identification of residues contributing to A2 domain-dependent structural stability in factor VIII and factor VIIIa. *J Biol Chem* 2008;283(17):11645-11651.
42. Bos MHA, Camire RM. Blood coagulation factors V and VIII: Molecular mechanisms of procofactor activation. *J Coagul Disord* 2010;2(2):19-27.
43. Venkateswarlu D. Role of hydrophobic mutations on the binding affinity and stability of blood coagulation factor VIIIa: A computational molecular dynamics and free-energy analysis. *Biochem Biophys Res Commun* 2014;450(1):735-740.
44. Gale AJ, Radtke KP, Cunningham MA, *et al*. Intrinsic stability and functional properties of disulfide bond-stabilized coagulation factor VIIIa variants. *J Thromb Haemost* 2006;4(6):1315-1322.
45. Lopez MJ, Mohiuddin SS. *Biochemistry, essential amino acids*. Treasure Island: StatPearls Publishing; 2026.