Received: 1 June 2020 | Revised: 14 October 2020 | Accepted: 3 November 2020

DOI: 10.1111/hae.14206

Haemophilia WILEY

ORIGINAL ARTICLE

Genetics

Mutation analysis in the *F8* gene in 485 families with haemophilia A and prenatal diagnosis in China

Yin Feng[1] | Qianqian Li[1] | Panlai Shi[1] | Ning Liu[1] | Xiangdong Kong[1] | Ruixia Guo[2]

[1]The Department of Obstetrics and Gynecology, The Genetics and Prenatal Diagnosis Center, The First Affiliated Hospital of Zhengzhou University, Henan, China

[2]The Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Henan, China

Correspondence
Ruixia Guo, The Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450002, Henan, China.
Email: grxcdxzzu@163.com

Xiangdong Kong, The Genetics and Prenatal Diagnosis Center, The Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450002, Henan, China.
Email: kongxd@263.net

Funding information
ZhongYuan thousand talents program-the Zhongyuan eminent doctor in Henan province, Grant/Award Number: ZYQR201810107; National Key R&D Program of China, Grant/Award Number: 2018YFC1002203

Abstract

Background: Haemophilia A (HA) is an X-linked bleeding disorder caused by mutations in the coagulation factor VIII (*F8*) gene. Its incidence in men is estimated to be approximately 1/5000.

Objective: This study aimed to characterize the mutation spectrum of the *F8* gene in 485 Chinese families, encompassing all HA phenotypic classes. Additionally, we evaluated the accuracy of prenatal diagnosis of foetuses at risk of having HA.

Methods: Long-Distance PCR (LD-PCR) and Multiplex PCR were used to detect inversions, next-generation sequencing (NGS) was used for point mutations, and multiplex ligation-dependent probe amplification (MLPA) was used for large deletions or duplications.

Results: A mutation spectrum of 478 HA families was produced. Throughout 26 exons and 15 introns, a total of 237 different alterations of mutations were detected, of which 146 are known mutations (64.5%) and 91 are novel mutations (35.5%). Prenatal diagnosis revealed 97 normal males (35.79%), 103 HA males (38.01%), 36 normal females (13.28%), and 38 HA carrier females (14.02%).

Conclusion: Using a systematic approach comprised of three steps, 237 pathogenic variants in 478 out of 485 patient samples (98.6%) were detected, including the identification of a heterogeneous mutation spectrum of 91 novel mutations. In addition, prenatal diagnosis of HA in pregnant carriers allowed for accurate determination of the foetal *F8* gene state.

KEYWORDS
F8, haemophilia A, mutation spectrum, prenatal diagnosis

1 | INTRODUCTION

Haemophilia A (HA) is the most common X-linked recessive bleeding disorder and is characterized as a deficiency or dysfunction of coagulation factor VIII. Patients with HA show symptoms of prolonged bleeding after injury or trauma. In the coagulation pathway, factor VIII and factor IX form an active complex (tenase complex) which activates subsequent factors, including factor X.[1] The prevalence of HA among live male births worldwide is approximately 1/5000.[2] Clinically, the phenotype of HA is classified based on residual factor VIII activity (FVIII:C), as severe (<1%), moderate (1%-5%), or mild (>5%).[3] The FVIII gene (*F8*), located at the distal end of the long arm of the X chromosome, spans 186 kb of genomic DNA and is comprised of 26 exons. The *F8* gene is translated into a 2315 amino acid

polypeptide chain with a 19 amino acid leader peptide. The mature polypeptide consists of six structural domains, namely A1-A2-B-A3-C1-C2, and is activated following the dissociation of the B domain.[4,5] In peripheral blood, the FVIII protein is bound to the von Willebrand factor (vWF), which protects it from degradation.

HA is characterized by extremely high mutational heterogeneity. To date, 3171 mutations of the *F8* gene identified as disease causing have been recorded in the HGMD® Professional 2020.1 database (https://portal.biobase-international.com/hgmd/). Overall, missense and nonsense variants are responsible for more than 50% of HA cases, followed by small deletions, duplications or insertions. However, inversions in intron 22 (Inv22) and intron 1 (Inv1) are the most frequent gene defects found in severe HA patients, occurring in 45–50% and 0.5–5% of cases, respectively.[6,7] It has been reported that both Inv22 and Inv1 may be correlated with the formation of inhibitors that interfere with treatment.[8]

The aim of this research is to produce a mutation spectrum for the *F8* gene based on 485 unrelated Chinese HA carrier families using a systematic approach, including LD-PCR, multiplex PCR, NGS and MLPA. Additionally, the accuracy of prenatal diagnosis of HA from the analysis of 274 foetal samples is reported.

2 | MATERIALS AND METHODS

2.1 | Participants

The study was approved by the ethics committee of the First Affiliated Hospital of Zhengzhou University. A total of 485 unrelated HA cases were included, comprising 381 severe, 72 moderate, and 32 mild clinical patients. The clinical phenotype was mainly classified based on factor VIII activity via biochemical characterization or analysis of bleeding characteristics. Once an informed consent document was signed, 2 mL of peripheral blood was collected in EDTA tubes from each participant.

2.2 | DNA extraction

Genomic DNA was obtained from 500 μL of whole blood using a Lab-Aid Nucleic Acid (DNA) Isolation Kit (Zeesan, Xiamen, China) according to the manufacturer's instructions. DNA purity and concentration was determined using UV spectrophotometry (Nanovue Plus, GE healthcare, Buckinghamshire, UK).

2.3 | Inversion detecting

Samples from severe cases were first tested for Inv22. The analysis was performed using an improved method based on the traditional Long-Distance PCR method,[9] which involves performing two PCR reactions concurrently in two distinct reaction tubes. The PCR products obtained were analysed using 0.7% agarose gel electrophoresis. Subsequently, samples identified as negative for Inv22 were subjected to multiplex PCR to evaluate Inv1, according to the procedure reported previously.[10] The PCR products obtained herein were analysed using 1% agarose gel electrophoresis.

2.4 | Next-generation sequencing (NGS)

DNA from inversion-negative patient samples with severe as well as moderate-to-mild HA were subjected to NGS of the essential regions of the *F8* gene, including all exons, both the 5'- and 3'-UTRs, and exon-intron junction regions. The entire procedure was conducted using the Life Technologies PGM platform (Carlsbad, CA, USA) according to the standard operating procedure.

2.5 | Multiplex ligation-dependent probe amplification (MLPA)

For patient samples lacking mutations, large deletions and duplications within the 26 exons of the *F8* gene were detected using MLPA. The process was carried out using the SALSA MLPA Probemix P178-A1 F8 Kit (MRC-Holland, Amsterdam, Netherlands) according to the manufacturer's instructions.

2.6 | Prenatal diagnosis

Two hundred and thirty-five pregnant probable HA carriers were confirmed to carry the familial mutation of the *F8* gene and applied for prenatal diagnosis. A total of 274 samples, including chorionic villi, direct amniocytes, or umbilical cord blood cells, were collected. Potential maternal contamination of foetal samples was identified using Goldeneye™ DNA ID System 20A Kit (PEOPLESPOT, Beijing, China). Subsequently, all foetal samples underwent the same genetic analysis as the probands.

2.7 | Data analysis

The data resulting from the PGM runs was analysed using Ion Torrent sequencing data analysis software from Torrent Suite (Life Technologies, Carlsbad, CA, USA) to generate sequence reads. Finally, all resulting gene variations were collected and analysed.

2.8 | Mutation verification and software prediction

Sanger sequencing was performed to verify the nucleotide sequence of all generated variants. To determine the pathogenicity of novel

identified missense variants, a variety of functional prediction software packages were used, including Polyphen_2[11] (http://genetics.bwh.harvard.edu/pph2/), PROVEAN[12] (http://provean.jcvi.org/index.php) and MutationTaster2[13] (http://www.mutationtaster.org/).

3 | RESULTS

Following the molecular diagnostic workflow described above, 237 different alterations of the *F8* gene were identified throughout all 26 exons and 15 introns, in 478 out of 485 (98.56%) unrelated HA patients. Upon comparison of the results with available databases, including dbSNP, Factor VIII Variant (http://www.factorviii-db.org/) and HGMD® Professional 2020.1, 146 known mutations and 91 novel mutations were identified. The distribution of mutational characteristics of the *F8* gene mutations identified is shown in the Table S1.

In 374 of the 381 severe cases, inversion tests identified 167 patients with Inv22 and 10 patients with Inv1, representing 44.65% and 2.67% of the severe cases evaluated, respectively. Frameshift mutations encompassed the second most common mutation type identified and accounted for 62 of the severe cases studied (16.58%). These mutations included 40 small deletions, 19 small insertions and t indels representing 10.70%, 5.08%, 0.80% of the severe cases evaluated, respectively. In addition, we observed 53 missense mutations (14.17%), 41 nonsense mutations (10.96%), 17 splice site mutations (4.55%), 16 gross deletions (4.28%), five gross insertions (1.34%) and three complex mutations (0.80%). Notably, a single female patient with a severe phenotype was found to carry an Inv1 mutation. Among the moderate cases, the dominant mutation was missense, which contributed 50 cases, representing 69.44% of all moderate phenotypes studied. Less common mutation types identified within this HA phenotype class included Inv22 (4.17%), small deletion (12.5%), splice site mutation (8.33%) and gross duplication (1.39%). Among patients with a mild phenotype, the mutation spectrum consisted of 29 cases of missense mutation (90.63%), one nonsense mutation (3.13%), one small insertion (3.13%) and one gross deletion (3.13%). The two female patients with mild HA symptoms were found to carry a missense mutation and a nonsense mutation. The data described above are summarized in Table 1.

TABLE 1 Frequencies of mutation type by severity

Mutation type	Severe N (% of total)	Moderate N (% of total)	Mild N (% of total)
Inv22	167 (44.65)	3 (4.17)	0
Inv1	10 (2.67)	0	0
Missense	53 (14.17)	50 (69.44)	29 (90.63)
Nonsense	41 (10.96)	3 (4.17)	1 (3.13)
Small deletion	40 (10.70)	9 (12.5)	0
Small insertion	19 (5.08)	0	1 (3.13)
Small indels	3 (0.80)	0	0
Splicing	17 (4.55)	6 (8.33)	0
Gross deletion	16 (4.28)	0	1 (3.13)
Gross duplication	5 (1.34)	1 (1.39)	0
Complex	3 (0.80)	0	0
Total	374 (100)	72 (100)	32 (100)

Note: Inv1: inversion in intron 1; Inv22: inversion in intron 22; Small deletion, small insertion, small indel: 20 bp or less.

Based on their clinical phenotype, 235 pregnant HA carriers provided 274 prenatal samples. The results of the prenatal diagnosis are presented in Table 2. Via sex identification, there were 200 males and 74 females. Of the 200 male foetuses, 103 were found to carry the familial HA mutation. Of the 74 female foetuses, 38 were carriers while the other 36 were mutation free.

4 | DISCUSSION

In this study, we used a variety of analytical methods to evaluate the HA mutation spectrum in Chinese families comprising all phenotypic classes. Specifically, LD-PCR and multiplex PCR was first used to detect Inv22 and Inv1, respectively, followed by NGS for point mutations, MLPA for deletion and duplication of exons, and Sanger sequencing for sequence verification. In total, 478 pathogenic variants in 485 HA patients were successfully identified with an efficiency of 98.56%. Despite the use of multiple techniques, no mutations were

TABLE 2 Consequence of prenatal diagnosis

Mutation type	Normal male foetus	Male patient	Normal female foetus	Female carrier	Total
Inv1	2	1	1	1	5
Inv22	29	43	6	18	96
Missense	32	25	11	10	78
Nonsense	10	9	3	3	25
Frameshift	15	16	12	5	48
Splicing	4	8	3	1	16
Gross deletion	5	1	0	0	6
Total	97 (35.40)	103 (37.59)	36 (13.14)	38 (13.87)	274

Note: Inv1: inversion in intron 1; Inv22: inversion in intron 22; Frameshift includes small deletion, small insertion and small indel.

identified in the remaining seven individuals. A possible explanation for this finding includes the presence of mutations located deep within introns or in regions outside of the *F8* gene that are important for its expression.[14]

Inv22 and Inv1 mutations in patients with severe HA have been reported in the ranges of 40–50% and 0.5–5%, respectively. In this study, the prevalence of Inv22 and Inv1 mutations were found to be similar to those described in other populations.[10,15] Intron 22 is the largest intron in the *F8* gene, and it has been established that its inversion impedes the formation of full-length FVIII messenger RNA (mRNA), thus precluding the synthesis of the FVIII protein, and resulting in severe HA.[16] Nonetheless, three cases of Inv22 mutations were detected in patients with moderate phenotypes. This could be due to clinical variability and gene variation among different individuals. A milder bleeding phenotype has been reported in men with severe haemophilia A. To date, there have been a few reports of a milder bleeding phenotype in patients with severe haemophilia A.[17,18] Researchers have suggested the following potential reasons for this observation: (i) DNA replication/RNA transcription errors resulting in restoration of the reading frame, (ii) 'ribosomal frameshifting', and (iii) *de novo* mutations with somatic mosaicism that occurred post-zygotically.[19,20] However, alternative causes, such as mutations in the anti-coagulation genes, may also be contributing factors. Similarly, a large deletion within the *F8* gene resulting in a mild phenotype could result from similar factors. Therefore, it is necessary to collect accurate clinical materials (e.g. FVIII:C, inhibitors, family history and arthropathy) and focus on phenotypic and genotypic data in genetic analysis. In addition, further studies addressing the incongruity between phenotype and genotype are necessary.

Among the severe cases, a female patient whose parents did not carry the familial HA mutation was confirmed to have the Inv1 mutation, suggesting that the mutation must have occurred *de novo* in the paternal germline. Due to the absence of a second mutation in this female patient, the severe symptoms of her might be associated with a nonrandom pattern of inactivation of the maternally derived X chromosome.[21] Of the two female patients with mild symptoms, one was identified as having a missense mutation, and the other a nonsense mutation. The father of the female with a missense mutation is also an HA patient, and his clinical presentation is moderate. The female, as the unique patient in her family, with a nonsense mutation was verified to have inherited the mutation from her mother. This single-nucleotide substitution is in the B domain, a region lacking procoagulant activity and is partially spliced from the mature protein.[22] Due to the combination of individual differences with respect to the mechanism of a nonrandom pattern of inactivation of the maternally derived X chromosome, the female presented mild clinical symptoms.

In this study, a total of 160 types of single-nucleotide substitutions were identified in 253 unrelated families, including 101 missense mutations in 132 HA families, 38 nonsense mutations in 45 HA families and 21 splice site mutations in 23 HA families. Of the 160 single-nucleotide substitutions identified, 51 were either not included in the available databases or not previously reported in the literature. All novel single-nucleotide substitutions were predicted to be damaged using computational analysis. Since some missense mutations may present polymorphisms,[23] further evaluation may be necessary. Moreover, a nonsense mutation, c.1063C>T, was identified in four unrelated families, as well as two missense mutations, c.2167G>A and c.6506G>A, that were detected in six and four unrelated families, respectively. These three mutations may be regarded as recurrent mutations for single-nucleotide substitutions.

Frameshift mutations were prominently observed in this study, and included small deletions, insertions and indels. All frameshift mutations detected, except two (c.599_601delAAG and c.7031delG), resulted in premature stop codons. Since exon 14 is the longest exon in the *F8* gene and possesses a large number of poly-A runs, it is considered a hot spot for such mutations.[24] Within a sequence of nine adenines in exon 14, a deletion (c.3637delA) was identified in 11 unrelated HA families and an insertion (c.3637insA) was detected in two unrelated HA families. Thus, the single-nucleotide deletion/insertion at this site represents a mutation hot spot within our cohort. Additional frameshift mutations observed include a three-base deletion (c.599_601delAAG mutation) causing an in-frame amino acid deletion (p.Glu200del) within the A2 structural domain, and a single base deletion at the end of exon 26 (c.7031delG), leading to the loss of the termination codon and a 39 amino acid extension to the polypeptide chain (p.Gly2344Alafs*39). Future analysis is required to elucidate the mechanistic details surrounding the impact these mutations have on FVIII protein expression. Finally, 22 large rearrangements were identified using MLPA technology, including 16 gross deletions spanning all exons, and six gross duplications spanning all exons, except exon 1. Two large deletions (c.4740+?_4933-?del, c.5054+?_5219-?del) between 150 and 200 bases in exon 14 were screened by NGS and verified by RT-PCR. Although their existence was carefully established, the accurate break points remain unknown. Further research is necessary to identify the exact break points and assess their pathogenicity.

In developed countries, replacement therapy for haemophilia is an effective treatment option, and prenatal diagnosis resulting in termination of pregnancy is not commonly practiced. Due to economic pressure and social attitudes in developing countries, prenatal diagnosis is urgently needed to prevent births of babies with haemophiliac phenotypes. As a result of the prenatal diagnosis carried out in this study, two women with prenatally diagnosed haemophiliac sons continued their pregnancy, while the other 101 women decided to terminate their pregnancies. A follow-up was carried out and the phenotypes of the newborns were consistent with the results of the prenatal diagnosis.

5 | CONCLUSION

This study revealed a mutation spectrum of the *F8* gene in a large sample of HA patients in China. Ninety-one novel mutations were reported, including 51 single-nucleotide substitutions, 33 frameshift mutations, two gross deletions, three gross insertions and two complex mutations. The mechanism by which the identified mutations affect disease severity requires additional research in order to provide

a better understanding of the genotype-phenotype association in HA. The mutation profiles of HA obtained here will provide a useful complementary database for the detection and diagnosis of patients and carriers of HA in China. In addition, the prenatal diagnosis of 274 samples from 235 pregnant HA carriers was determined and verified in two cases where the pregnancies were carried to term. Our results demonstrate that potential HA carriers should define their status prior to pregnancy, so that timely prenatal diagnosis, if desired, is possible.

ACKNOWLEDGEMENTS

This work was supported by the ZhongYuan thousand talents program-the Zhongyuan eminent doctor in Henan province (No. ZYQR201810107) and National Key R&D Program of China (No. 2018YFC1002203).

CONFLICT OF INTERESTS

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

AUTHOR CONTRIBUTIONS

YF performed experiments, analysed the data, prepared tables and figures and wrote the manuscript. Q-Q L, P-L S and NL interpreted the data, edited the manuscript. R-X G and X-D K designed the study, supervised the interpretation and statistical analysis of the data, edited the manuscript. All the authors approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available within the article and its supplementary information files.

ORCID

Yin Feng https://orcid.org/0000-0002-1598-3156

REFERENCES

1. Bowen DJ. Haemophilia A and haemophilia B: molecular insights. *Mol Pathol*. 2002;55(2):127-144.
2. Bolton-Maggs PH, Pasi KJ. Haemophilias A and B. *Lancet*. 2003;361(9371):1801-1809.
3. Keeney S, Mitchell M, Goodeve A. The molecular analysis of haemophilia A: a guideline from the UK haemophilia centre doctors' organization haemophilia genetics laboratory network. *Haemophilia*. 2005;11(4):387-397.
4. Gitschier J, Wood WI, Goralka TM, et al. Characterization of the human factor VIII gene. *Nature*. 1984;312(5992):326-330.
5. Woods-Samuels P, Kazazian HH Jr, Antonarakis SE. Nonhomologous recombination in the human genome: deletions in the human factor VIII gene. *Genomics*. 1991;10(1):94-101.
6. Albánez S, Ruiz-sáez A, Boadas A, De bosch N, Porco A. Identification of factor VIII gene mutations in patients with severe haemophilia A in Venezuela: identification of seven novel mutations. *Haemophilia*. 2011;17(5):e913-e918.
7. Cumming AM. The factor VIII gene intron 1 inversion mutation: prevalence in severe hemophilia A patients in the UK. *J Thromb Haemost*. 2004;2(1):205-206.
8. Castaldo G, D'Argenio V, Nardiello P, et al. Haemophilia A: molecular insights. *Clin Chem Lab Med*. 2007;45(4):450-461.
9. Liu Q, Nozari G, Sommer SS. Single-tube polymerase chain reaction for rapid diagnosis of the inversion hotspot of mutation in hemophilia A. *Blood*. 1998;92(4):1458-1459.
10. Bagnall RD, Waseem N, Green PM, Giannelli F. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood*. 2002;99(1):168-174.
11. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.
12. 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.
13. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 2014;11(4):361-362.
14. El-Maarri O, Singer H, Klein C. Lack of F8 mRNA: a novel mechanism leading to hemophilia A. *Blood*. 2006;107(7):2759-2765.
15. Mantilla-Capacho JM, Beltrán-Miranda CP, Luna-Záizar H, et al. Frequency of intron 1 and 22 inversions of Factor VIII gene in Mexican patients with severe hemophilia A. *Am J Hematol*. 2007;82(4):283-287.
16. Antonarakis SE, Rossiter JP, Young M, et al. Factor VIII gene inversions in severe hemophilia A: results of an international consortium study. *Blood*. 1995;86(6):2206-2212.
17. Miese F, Schleich C, Nebelung S, et al. Motion correction improves image quality of dGEMRIC in finger joints. *Eur J Radiol*. 2011;80(3):e427-e431.
18. Olsson A, Ljung R, Hellgren M, Berntorp E, Baghaei F. Phenotype and genotype comparisons in carriers of haemophilia A. *Haemophilia*. 2016;22(3):e235-e237.
19. Young M, Inaba H, Hoyer LW, Higuchi M, Kazazian HH, Antonarakis SE. Partial correction of a severe molecular defect in hemophilia A, because of errors during expression of the factor VIII gene. *Am J Hum Genet*. 1997;60(3):565-573.
20. Yenchitsomanus P, Akkarapatumwong V, Pung-Amritt P, et al. Genotype and phenotype of haemophilia A in Thai patients. *Haemophilia*. 2003;9(2):179-186.
21. Graw J, Brackmann H-H, Oldenburg J, Schneppenheim R, Spannagl M, Schwaab R. Haemophilia A: from mutation analysis to new therapies. *Nat Rev Genet*. 2005;6(6):488-501.
22. Thompson AR. Structure and function of the factor VIII gene and protein. *Semin Thromb Hemost*. 2003;29(1):11-22.
23. Ogata K, Selvaraj SR, Miao HZ, Pipe SW. Most factor VIII B domain missense mutations are unlikely to be causative mutations for severe hemophilia A: implications for genotyping. *J Thromb Haemost*. 2011;9(6):1183-1190.
24. Nakaya S, Liu ML, Thompson AR. Some factor VIII exon 14 frameshift mutations cause moderately severe haemophilia A. *Br J Haematol*. 2001;115(4):977-982.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Feng Y, Li Q, Shi P, Liu N, Kong X, Guo R. Mutation analysis in the *F8* gene in 485 families with haemophilia A and prenatal diagnosis in China. *Haemophilia*. 2021;27:e88–e92. https://doi.org/10.1111/hae.14206