scientific reports

Check for updates

OPEN

Xq28 duplication not F8 inversion: integrated genetic reanalysis redefines prenatal carrier diagnosis

Xueting Yang[1], Subinuer Maimaiti[2], Qingwei Qi[1], Xiya Zhou[1], Na Hao[1], Jiazhen Chang[1], Mengmeng Li[1], Kaili Yin[1], Yan Lü[1]✉ & Yulin Jiang[1]✉

This study aimed to resolve diagnostic discrepancies in which expanded carrier screening (ECS) in a healthy population suggested hemophilia A carrier status due to *F8* intron 22 inversion (Inv22), while further analysis revealed a true distal Xq28 duplication. We evaluated the clinical utility of integrated genetic technologies in distinguishing between these two entities. A healthy pregnant woman underwent ECS, with long-distance polymerase chain reaction (LD-PCR) specifically employed for *F8* Inv22 detection. Given the identified heterozygote of *F8* inversion, prenatal diagnosis was conducted, integrating LD-PCR, single-nucleotide polymorphism array (SNP array), optical genome mapping (OGM), third-generation long-read sequencing, and X-chromosome inactivation (XCI) analysis to characterize the variant thoroughly. Initial LD-PCR suggested *F8* Inv22 in both the fetus and the mother. SNP array revealed a 0.5-Mb Xq28 duplication spanning *F8* exons 1–22. OGM and long-read sequencing confirmed the occurrence of duplication while excluding the *F8* inversion, revealing *int22h-1/int22h-2*-mediated nonallelic homologous recombination as the mechanism. XCI analysis revealed skewed inactivation (18.1%) of the duplicated allele in the asymptomatic mother, whereas the fetus exhibited a random XCI pattern. This study highlights a critical diagnostic pitfall: even well-established techniques such as LD-PCR for *F8* Inv22 may yield misleading results when applied to general population screening beyond their original diagnostic context. This underscores the importance of accurate interpretation of ECS results through appropriate analytic validation. We propose a tiered strategy: LD-PCR screening followed by OGM or long-read sequencing to differentiate *F8* inversions from Xq28 duplications in the general population.

Keywords Distal Xq28 duplication syndrome, *F8* inversion, Prenatal diagnosis, Expanded carrier screening, Optical genome mapping, Long-read sequencing

Hemophilia A (HA), an X-linked coagulation disorder caused by pathogenic variants in the *F8* gene, is characterized by a deficiency in the production of coagulation factor VIII. Epidemiological studies have reported a prevalence of 24.6 per 100,000 male births, with severe forms (factor VIII activity < 1%) occurring in 9.5 per 100,000 male births[1]. Severe HA is characterized by frequent bleeding episodes with early onset before the age of two, leading to disabling hemophilic arthropathy and potentially life-threatening intracranial hemorrhage[2].

Expanded carrier screening (ECS) is performed during early pregnancy and preconception to inform the participants of possible genetic disease risks in future offspring and of the reproductive options available. ECS panels typically incorporate conditions meeting the criteria for moderate-to-profound phenotypes, early disease onset, and high carrier frequency. In accordance with these parameters, the American College of Medical Genetics and Genomics recommends the inclusion of HA in ECS panels[3].

The most common pathogenic variant in HA is *F8* inversion. Intron 22 inversion (Inv22) originates from nonallelic homologous recombination (NAHR) between *int22h*-1, a 9.5 kb segmental duplication element within intron 22, and two homologous duplication elements (*int22h*-2/*int22h*-3), located on the telomeric side of X[4,5]. Intron 1 inversion (Inv1) shares a similar molecular mechanism[6,7]. These two recurrent inversions account for most severe HA cases, making *F8* inversion-based testing a robust and efficient ECS strategy[7,8].

Notably, the *F8* gene partially resides within the distal Xq28 region flanked by *int22h*-1 and *int22h*-2. NAHRs between these two elements can also cause copy number changes. Xq28 deletions mediated by *int22h-1/int22h-2*

[1]Department of Obstetrics and Gynecology, National Clinical Research Center for Obstetric & Gynecologic Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science & Peking Union Medical College, Beijing, China. [2]Urumqi Maternal and Child Health Care Hospital, Urumqi, China. ✉email: lvyan@pumch.cn; yulinj@gmail.com

have been proposed to potentially contribute to embryonic lethality in males[9], a hypothesis that may account for the predominance of duplication cases observed in clinical settings. Distal Xq28 duplication may lead to cognitive impairment, neurobehavioral issues, and facial dysmorphism in males, while most female heterozygotes present mild or no symptoms[10].

The overlapping architecture of *F8* and Xq28 complicates genotype–phenotype predictions, as NAHR variants can cause HA, distal Xq28 duplication syndrome, or both. We report a clinically unaffected female identified as a potential heterozygote for *F8* Inv22 through ECS during early pregnancy. Further integrated genetic analysis revealed an Xq28 duplication without *F8* inversion in both the proband and the fetus. We investigated the mechanistic origins of the discordant testing results and emphasized caution when applying ECS for HA screening in healthy populations without a family history, even using a classical testing method.

Materials and methods

Study participants and initial screening

A 27-year-old healthy primigravida underwent ECS during early pregnancy. Our ECS panel covers genes related to 197 autosomal recessive and 26 X-linked diseases[11]. The ECS platform integrates next-generation sequencing with long-distance polymerase chain reaction (LD-PCR) for *F8* inversion detection and PCR for *FMR1* variant detection. The result indicated the presence of *F8* Inv22, suggesting the carrier status for HA. The primigravida reported no personal history of coagulation abnormalities and no familial history of bleeding disorders.

Prenatal genetic diagnosis

To confirm fetal status, amniocentesis was performed at 18^{+3}weeks to detect *F8* Inv22 through LD-PCR. Concurrently, karyotyping and single-nucleotide polymorphism array (SNP array) analyses were conducted to screen for aneuploidy and copy number variation (CNV). Driven by strong maternal concern, trio exome sequencing (ES) was also performed to exclude monogenetic disorders. Following the detection of both *F8* Inv22 and Xq28 duplication, which are located within the same genomic region, optical genome mapping (OGM) was subsequently utilized for structural characterization. As OGM yielded results inconsistent with prior findings, validation was performed using third-generation long-read sequencing. Furthermore, to assess potential phenotypic influences, X-chromosome inactivation (XCI) patterns in both the mother and fetus were investigated.

DNA extraction

Genomic DNA was extracted from amniotic fluid and EDTA-anticoagulant peripheral blood from the pregnant woman and her husband using a QIAamp DNA Mini Kit (Qiagen, Germany). Ultrahigh-molecular-weight (UHMW) DNA (>150 kb) was extracted using the Bionano Prep™ Blood and Cell Culture DNA Isolation Kit according to the manufacturer's instructions (Bionano Genomics, San Diego, CA, USA).

F8 Inv22 detection

Genotyping of Inv22 was performed using LD-PCR as described previously with some modifications[12]. The primers used were as follows: P, 5′- GCC CTG CCT GTC CAT TAC ACT GAT GAC ATT ATG CTG AC -3′; Q, 5′- GGC CCT ACA ACC ATT CTG CCT TTC ACT TTC AGT GCA ATA -3′; B, 5′- CCC CAA ACT ATA ACC AGC ACC TTG AAC TTA CCC TCT -3′. Briefly, the PCRs were performed in two separate tubes. In one tube, the amplification reaction was mediated by primers P and Q; in the other tube, the amplification was mediated by primers P and B. The existence and sizes of the LD-PCR products were confirmed using agarose gel electrophoresis. Negative and positive controls were used to ensure the accuracy of the sample results. LD-PCR was also used in ECS for the detection of *F8* Inv22 and Inv1.

Xq28 duplication detection

SNP array utilizing the Affymetrix CytoScan 750 K platform was routinely employed to detect CNVs. Genome coordinates were referenced according to the hg19/GRCh37 human reference genome assembly.

Distinguishing Xq28 duplication from F8 Inv22

OGM has been shown to have advantages in detecting structural variants and CNVs based on UHMW DNA[13]. The extracted UHMW DNA was fluorescently labeled with the enzyme DLE-1 using a DLS DNA Labeling Kit (Bionano Genomics, San Diego, USA). Then, the labeled DNA was loaded onto a Saphyr chip for linearization and imaging on the Saphyr instrument. The de novo assembly was executed with Bionano Solve software v.3.7, and OGM-specific pipelines were managed with Bionano Access software v.1.7. Variants were analyzed according to the institute's standard operating criteria.

Single molecule real-time (SMRT) sequencing is a third-generation long-read sequencing technology. Amplification of the *F8* gene and PacBio sequencing were performed as previously described[14,15]. Briefly, multiplex long-range PCR was performed with five primers (H1F, H1R, H2F, H3F, and H2/3R) to generate specific amplificons indicating *int22h*-1, *int22h*-2, and *int22h*-3, as well as *int22h*-mediated inversions, deletions, and duplications[5]. After purification and end repair, double barcode adapters were ligated to the 5′ and 3′ ends, and SMRTbell libraries were prepared using the Sequel Binding and Internal Ctrl Kit 3.0 (Pacific Biosciences, Menlo Park, CA). Subsequently, sequencing was initiated on the PacBio Sequel II System to generate 10 to 25 subreads per molecule, and the consensus circular sequence was mapped to the GRCh38 reference. Variants were called after alignment using FreeBayes software version 1.2.0 (https://www.geneious.com/plugins/freebayes; Biomatters, Inc., San Diego, CA).

Other genetic analyses

Karyotyping was routinely performed for the detection of aneuploidy and chromosomal structural abnormalities. The harvesting of cells, preparation of slides, staining with Giemsa and subsequent analysis were conducted according to the institute's standard operating procedures, and chromosomal abnormalities were described in accordance with the International System for Human Cytogenetic Nomenclature.

Trio-ES was employed for the detection of single nucleotide variants (SNVs). An Agilent liquid capture system (Agilent SureSelect Human All Exon V6, Agilent Technologies, CA, USA) was used to enrich exome sequences for library preparation in ES. The DNA libraries were sequenced on an Illumina platform, and 150 bp paired-end reads were generated. The obtained reads were mapped to the GRCh37/hg19 reference genome database after quality control. The Integrative Genomics Viewer (IGV) was used for sequence data visualization.

To analyze XCI at the DNA methylation level, genomic DNA was digested both with and without the methylation-sensitive restriction enzyme *HpaII*. DNA from each sample was subsequently subjected to PCR amplification of the *AR* (androgen receptor) gene and polymorphic loci on Xp (XpA, XpB, and XpC). The PCR products were resolved by capillary electrophoresis on an ABI 3500 Analyzer and quantified through peak height analysis using GeneMapper software. X-inactivation ratios were derived by calculating the ratio of allele products from digested DNA to those from undigested DNA, expressed as percentages.

Post-test counseling

Comprehensive post-test counseling was delivered to the couple, integrating all genetic findings to address the phenotypic implications, available treatments, long-term prognosis, inheritance mode, and recurrence risk for any future pregnancy related to the disease.

Results

Concurrent detection of inversion and duplication

LD-PCR suggested the presence of *F8* Inv22 in both the fetus and the mother (Fig. 1). Karyotyping revealed no chromosomal abnormalities. SNP array revealed a 514-kb duplication at the distal Xq28 region in the female fetus {arr[GRCh37] Xq28 (154121324_154625570) × 3}. At the genomic level, the CNV contains the *int22h*1/*int22h*2-flanked Xq28 recurrent region, fulfilling pathogenic classification criteria[16]. At the gene level, the CNV spans exons 1–22 of the *F8* gene and seven other OMIM genes, including *RAB39B* and *CLIC2*. SNP array also established the maternal origin of this CNV through identification of a corresponding 445-kb duplication in the mother (Fig. 2A). Trio-ES corroborated these findings, identifying a 454-kb Xq28 duplication in both the fetus and the mother {seq[GRCh37] Xq28(154124531_154578447) × 3} (Fig. 2B). No other pathogenic genetic variants were detected. Figure 3 integrates the genomic coordinates of the detected duplication with the classic *int22h1/int22h2*-flanked Xq28 recurrent region.

Exclusion of *F8* Inv22

Given the possible co-occurrence of inversion and duplication in the same genomic region and the detection limitations in LD-PCR, we employed OGM to elucidate the complex variation. While OGM successfully confirmed the ~ 0.5 Mb duplication in both subjects {ogm[GRCh37] dup(X)(q28q28) chrX:154,105,504–154,596,148}, it failed to detect the previously suspected *F8* Inv22.

To address the discrepancy between LD-PCR and OGM results, we applied SMRT sequencing for validation. The long-read data confirmed an *int22h*-mediated duplication and excluded *F8* inversion.

Mechanistic origins of the discordant testing results

The homologous duplication elements in *int22h*-1 and *int22h*-2 prompted NAHR, leading to tandem duplication of the intervening sequences. Notably, this recombination event did not induce *F8* inversion. The direct orientation of this tandem duplication, along with the single breakpoint within the *F8* gene, ensures the preservation of genomic integrity, thereby implying the maintenance of *F8* functionality.

In the LD-PCR analysis, primers P and Q are positioned within the *F8* gene at positions + 1334 bp and − 1212 bp flanking *int22h*-1, respectively, whereas primer B is situated − 118 bp upstream of *int22h*-2 and + 118 bp downstream of *int22h*-3. The wild-type X chromosome exclusively contains 12-kb PQ amplicons. In contrast, the Inv22 variant produces only 11-kb PB amplicons. Notably, the duplication variant mediated by *int22h*-1/*int22h*-2 NAHR results in both PQ and PB amplicons. According to established LD-PCR interpretation criteria, detection of the PB amplicons serves as an indicator of *F8* Inv22. Based on these findings, both the female fetus and the mother were identified as potential Inv22 heterozygotes.

As shown in Fig. 2C, OGM provided direct visualization of the genomic duplication event in the absence of inversion.

In SMRT sequencing analysis, *int22h*-related rearrangements mediated by different primer combinations generate specific amplicon profiles that distinguish wild-type X chromosomes, Inv22, duplication variants, and others. The combined set of H1F + H2/3R and H1R + H3F amplicons signifies Inv22 type 1, and the set of H1F + H2/3R and H1R + H2F amplicons signifies Inv22 type 2, whereas the presence of solo H1F + H2/3R amplicons indicates the duplication variant[5]. As shown in Fig. 2D, the wild-type X chromosome is characterized by a combined set of three distinct amplicons: H1F + H1R, H2F + H2/3R, and H3F + H2/3R. The additional duplication amplicons of H1F + H2/3R and their ratio to the combined set of wild-type X chromosome copies provide critical evidence for diagnosing the carrier status of Xq28 duplication without inversion.

XCI pattern

As shown in Table S1, the XCI patterns of the mothers and fetuses were determined. Assays revealed that the preferentially active allele in the mother corresponded to the allele transmitted to the fetus. This suggests that

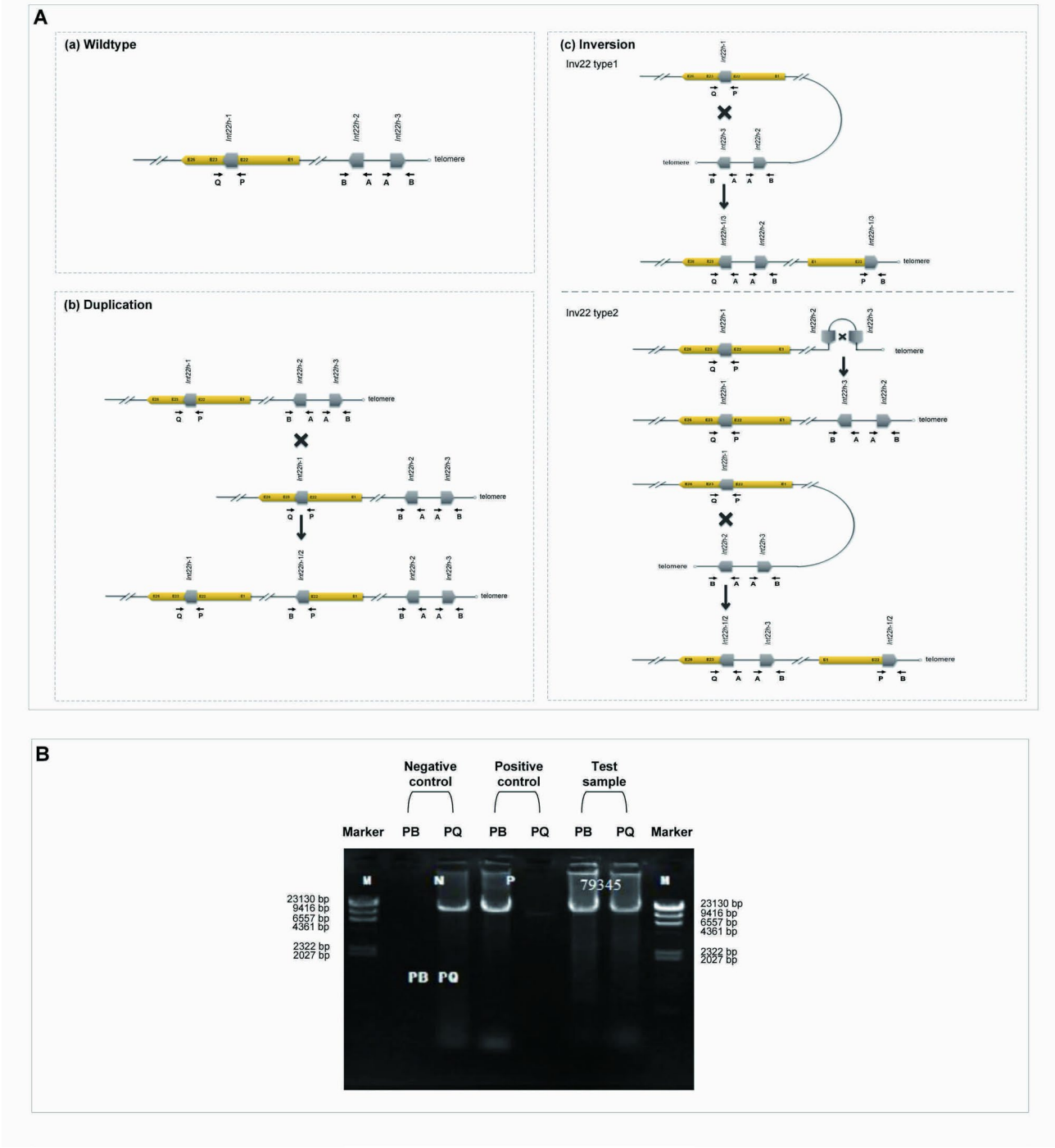

Fig. 1. Nonallelic homologous recombination (NAHR) events among the *int22h* sequences and long-distance polymerase chain reaction (LD-PCR) testing for the *int22h*-related inversion. (**A**) Schematic mechanism of the recurrent *int22h*-related rearrangements, (**a**) wild-type status of *F8* and *int22h* sequences in the Xq28 region; (**b**) duplication mediated by NAHR between *int22h*-1 and *int22h*-2; (**c**) inversions mediated by recombination among *int22h* copies on the same allele: Inv22 type 1 results from recombination of *int22h*-1 and *int22h*-3, whereas Inv22 type 2 is mediated by recombination of *int22h*-1 and *int22h*-2 (*int22h*-2 and *int22h*-3 first undergo intra-chromosomal recombination mediated by palindrome sequences, resulting in inverted orientations of *int22h*-1 and *int22h*-2). (**B**) LD-PCR followed by agarose gel electrophoresis for the amino fluid sample numbered 79,345. PB: amplicons of primer P + primer B; PQ: amplicons of primer P + primer Q.

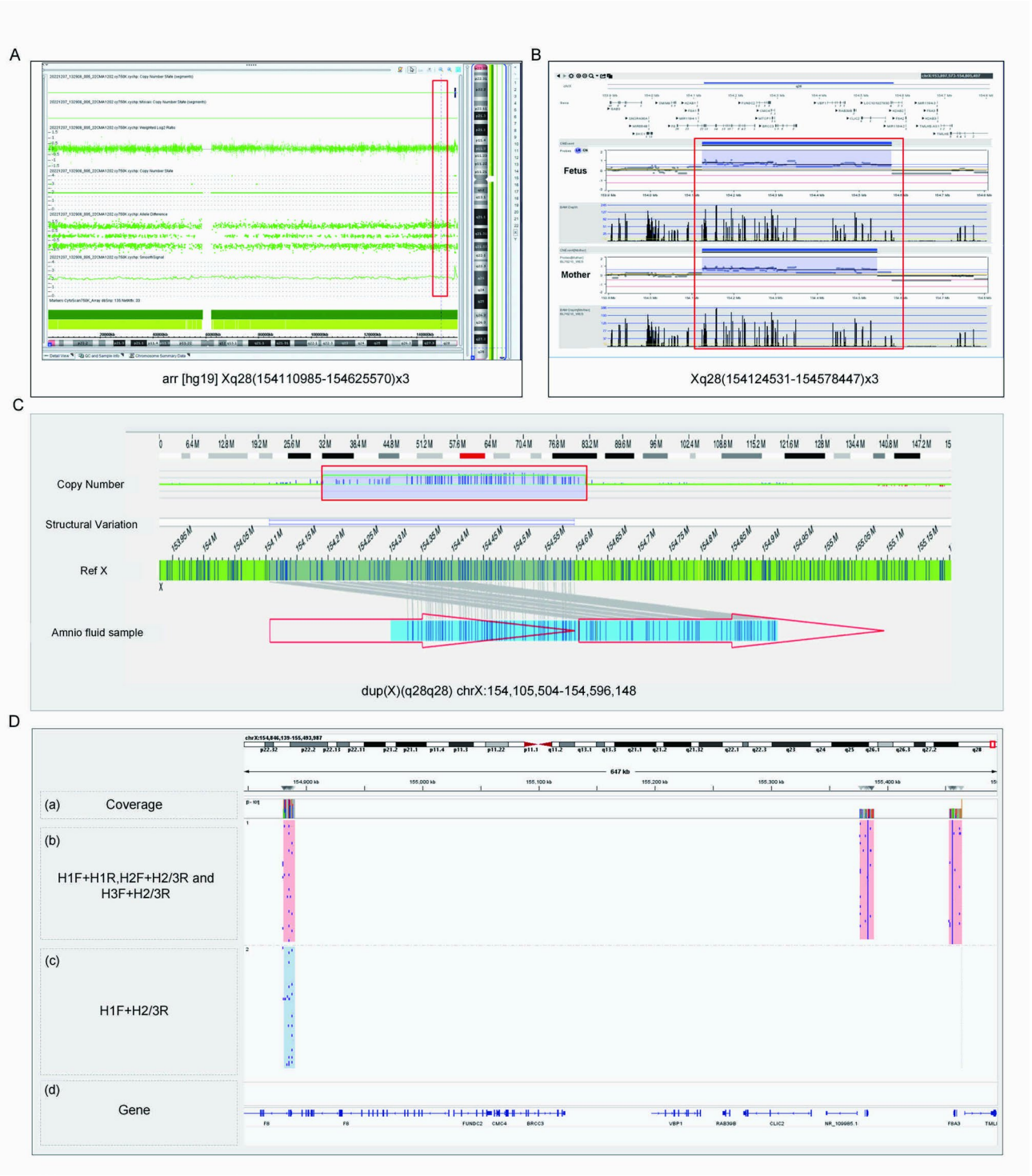

Fig. 2. Identification of the duplication and exclusion of inversion. (**A**) SNP array revealing the duplication in Xq28 (highlighted in the red frame). (**B**) Trio-ES reveals the fetus carrying the duplication inherited from the mother. (**C**) OGM reveals the exact location and orientation of the duplication. (**D**) IGV plots showing the SMRT sequencing results of amplicons with five primers (H1F, H1R, H2F, H3F, and H2/3R): (**a**) coverage of the sequencing reads shows the copy ratios of different amplicons; (**b**) amplicons representing copies of *int22h*-1, *int22h*-2 and *int22h*-3; (**c**) amplicons representing the duplication mediated by *int22h*-1 and *int22h*-2; (**d**) genes included in this region.

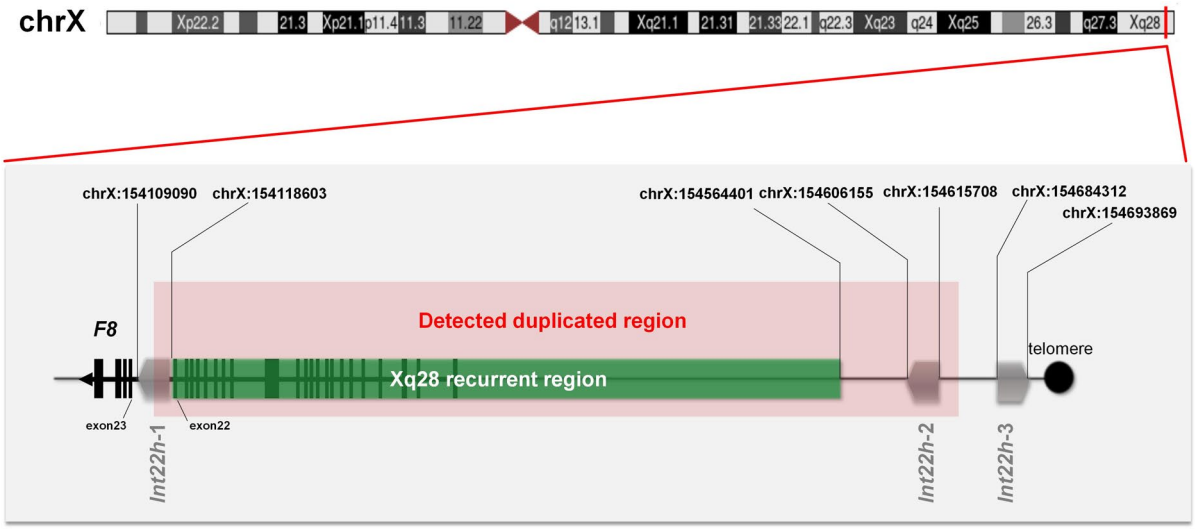

Fig. 3. Scheme of the location relationships among the *F8* gene, the detected duplication, the distal Xq28 recurrent region, and the *int22h* sequences. Exons of *F8* are shown by black rectangles; different *int22h* sequences are displayed in dark gray; the Xq28 recurrent region is shown in a green shallow; the duplicated region detected in this case is shown in a red shallow; and the genome reference of the locations is GRCh37hg19.

the X chromosome carrying the Xq28 duplication was preferentially active in the mother, with an average XCI ratio of 18.1%. In contrast, the fetus exhibited a random XCI pattern, with an average XCI ratio of 65.5% for the maternally inherited allele.

Post-testing counseling

The mother and her female fetus were diagnosed as heterozygotes for distal Xq28 duplication. The mother reported holding a college degree and formal employment. While most heterozygous females exhibit no apparent health or learning impairments and are clinically unaffected, a minority present with mild intellectual disability. Although the random XCI pattern observed in the fetus may suggest an unaffected status, the phenotypic inconsistency with the mother's skewed XCI pattern complicates precise phenotypic prediction. Given the 50% recurrence risk, preimplantation genetic testing is available for future pregnancies.

Discussion

In this study, we report a case of a healthy pregnant woman who underwent prenatal diagnosis due to the possible carrier status of HA identified by ECS utilizing LD-PCR for *F8* inversion detection. However, integrated genetic analyses combining SNP array, trio-ES, OGM, and long-read sequencing revealed the presence of a pathogenic duplication associated with distal Xq28 duplication syndrome, whereas no *F8* inversion was detected.

Among the current methodologies for *F8* inversion detection, LD-PCR remains the most widely implemented method because of its rapid turnaround time, cost-effectiveness, and diagnostic reliability[17,18]. Although Xq28 duplication heterozygotes show quantitatively elevated PQ/PB ratios compared with *F8* Inv22 heterozygotes as shown in Table S2, LD-PCR solely cannot resolve these dosage discrepancies, thereby precluding differentiation between these genotypes[19,20]. In PB-positive cases, CMA can readily distinguish isolated F8 inv22 from Xq28 duplications based on the absence or presence of a duplication. However, CMA cannot determine whether the duplication occurs in isolation or concurrently with inv22, as both scenarios yield identical results in LD-PCR (using PB + PQ amplicons). Cases with co-occurring variants have been documented previously[21]. This distinction is clinically important due to their differing phenotypic implications. Long-read sequencing or OGM—both long-molecule technologies—are required to resolve both variants simultaneously.

Long-read sequencing detects diverse HA-associated variants, including SNVs, indels, CNVs, and structural rearrangements; however, its clinical implementation is constrained by complex bioinformatics and stringent sample requirements[22,23]. OGM provides precise characterization of variant location, orientation, and breakpoints for CNVs and structural rearrangements within a single assay[24–26]. The substantially higher costs and extended turnaround times of these two methods justify a tiered diagnostic strategy: initial screening via LD-PCR followed by targeted long-molecule analysis for PB-positive cases. This approach balances efficiency and diagnostic precision while optimizing resource allocation. From a health-economic perspective, the benefit–cost ratio of ECS for *F8* Inv22 is approximately 2.1 (Supplementary File 3). Notably, because the screening cost reflects the entire gene panel rather than Inv22 alone, the cost-effectiveness of ECS for preventing other monogenic disorders may be even more favorable.

The association between *F8* gene duplications and HA manifestations depends on variant localization, exon length, and whether the resulting protein is in-frame or out-of-frame[27,28]. In our study, the mother and fetus were heterozygous for the classical ~ 0.5 Mb Xq28 duplication mediated by *int22h*1/2 NAHR, which is typically

not at risk of HA because it preserves functional *F8* activity. However, HA symptoms may occur if a secondary pathogenic variant, such as a disruptive SNV in the preserved *F8* allele, arises[29]. In our study, ES and OGM precluded such genetic abnormalities.

For distal Xq28 duplication syndrome, current evidence implicates dosage alterations in two candidate genes, *CLIC2* and *RAB39B*, as the probable molecular basis for intellectual disability[30]. Although most female heterozygotes are clinically unaffected, some exhibit neurocognitive deficits[10,31]. The association between skewed XCI and phenotypic manifestations in female heterozygotes remains unclear, with conflicting evidence in the literature[9,10]. Two key challenges in XCI testing are (1) the unknown threshold for clinical manifestation and (2) possible tissue-specific XCI mosaicism. In our study, the skewed inactivated normal allele was found in asymptomatic mothers, highlighting the limitations associated with predicting clinical phenotypes on the basis solely of XCI patterns. Moreover, recent data have identified five unaffected male carriers across four unrelated families, revealing previously unrecognized incomplete penetrance. These findings indicate that the duplication is not invariably associated with a syndromic phenotype and point to potential modifying mechanisms—such as a second genetic hit—that require further investigation. Accordingly, prenatal interpretation of Xq28 duplications should be approached with caution[32].

The limitations of this study are, first, its reliance on a single case; therefore, the proposed diagnostic strategy—integrating LD-PCR screening with OGM or long-read sequencing confirmation—requires further validation through accumulation of additional cases. Additionally, the absence of XCI skewing in the fetus does not guarantee an unaffected phenotype; however, loss to follow-up of pregnancy outcomes precluded phenotypic assessment of the offspring.

Conclusion

Our study serves as a compelling reminder that even well-established diagnostic methods can produce ambiguous outcomes when deployed beyond their validated scope, as evidenced by the misinterpretation of an Xq28 duplication as an *F8* inversion in ECS. Our integrated analysis successfully delineates the intricate rearrangements in this region and underscores the critical need for confirmatory testing. Consequently, we propose a tiered diagnostic pathway—initial LD-PCR screening coupled with definitive long-molecule verification—which translates this critical insight into a practical clinical strategy.

Data availability

The datasets generated and/or analysed during the current study are available in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2025) in National Genomics Data Center (Nucleic Acids Res 2025), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA014988) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human[33,34].

Received: 17 November 2025; Accepted: 5 March 2026
Published online: 25 March 2026

References

1. Iorio, A. et al. Establishing the prevalence and prevalence at birth of hemophilia in males: A meta-analytic approach using national registries.. *Ann. Intern. Med.* **171**, 540–546 (2019).
2. Mannucci, P. M. & Franchini, M. Haematology clinic: Haemophilia A.. *Hematology* **19**, 181–182 (2014).
3. Gregg, A. R. et al. Screening for autosomal recessive and X-linked conditions during pregnancy and preconception: A practice resource of the American College of Medical Genetics and Genomics (ACMG).. *Genet. Med.* **23**, 1793–1806 (2021).
4. Bagnall, R. D., Giannelli, F. & Green, P. M. Polymorphism and hemophilia A causing inversions in distal Xq28: A complex picture.. *J. Thromb. Haemost.* **3**, 2598–2599 (2005).
5. Bagnall, R. D., Giannelli, F. & Green, P. M. Int22h-related inversions causing hemophilia A: A novel insight into their origin and a new more discriminant PCR test for their detection.. *J. Thromb. Haemost.* **4**, 591–598 (2006).
6. Brinke, A. et al. Two chimaeric transcription units result from an inversion breaking intron 1 of the factor VIII gene and a region reportedly affected by reciprocal translocations in T-cell leukaemia.. *Hum. Mol. Genet.* **5**, 1945–1951 (1996).
7. Bagnall, R. D., Waseem, N., Green, P. M. & Giannelli, F. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A.. *Blood* **99**, 168–174 (2002).
8. Antonarakis, S. E. et al. Factor VIII gene inversions in severe hemophilia A: Results of an international consortium study.. *Blood* **86**, 2206–2212 (1995).
9. El-Hattab, A. W. et al. Int22h-1/int22h-2-mediated Xq28 rearrangements: Intellectual disability associated with duplications and in utero male lethality with deletions.. *J. Med. Genet.* **48**, 840–850 (2011).
10. El-Hattab, A. W. et al. Clinical characterization of int22h1/int22h2-mediated Xq28 duplication/deletion: New cases and literature review.. *BMC Med. Genet.* **16**, 12 (2015).
11. Hou, W. et al. Carrier screening for 223 monogenic diseases in Chinese population: A multi-center study in 33 104 individuals. *Nan Fang Yi Ke Da Xue Xue Bao* **44**, 1015–1023 (2024).
12. Liu, Q., Nozari, G. & Sommer, S. S. Single-tube polymerase chain reaction for rapid diagnosis of the inversion hotspot of mutation in Hemophilia A. *Blood* **92**, 1458–1459 (1998).
13. Yin, K. et al. Optical genome mapping to decipher the chromosomal aberrations in families seeking for preconception genetic counseling. *Sci Rep.* **15**, 2614 (2025).
14. Liang, Q. et al. A more universal approach to comprehensive analysis of Thalassemia alleles (CATSA). *J. Mol. Diagn.* **23**, 1195–1204 (2021).
15. Liu, Y. et al. Comprehensive analysis of Hemophilia A (CAHEA): Towards full characterization of the F8 gene variants by long-read sequencing. *Thromb. Haemost.* **123**, 1151–1164 (2023).
16. Riggs, E. R. et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: A joint consensus recommendation of the American college of medical genetics and genomics (ACMG) and the clinical genome resource (ClinGen). *Genet. Med.* **22**, 245–257 (2020).
17. González-Ramos, I. A. et al. Genetic analysis for carrier diagnosis in Hemophilia A and B in the Mexican population: 25 Years of experience. *Am. J. Med. Genet. C Semin. Med. Genet.* **184**, 939–954 (2020).

18. Mahmoud Abu Arra, C., Samarah, F. & Sudqi Abu Hasan, N. Factor VIII intron 22 inversion in severe hemophilia A patients in Palestine. *Scientifica* **2020**, 3428648 (2020).
19. Green, P., Hill, M. & Bowen, D. J. Duplications involving int22h-1 of the factor VIII gene: A cause for concern in genetic testing for Hemophilia A?. *J. Thromb. Haemost.* **5**, 2155–2156 (2007).
20. Rossetti, L. C., Radic, C. P., Larripa, I. B. & De Brasi, C. D. Developing a new generation of tests for genotyping hemophilia-causative rearrangements involving int22h and int1h hotspots in the factor VIII gene. *J. Thromb. Haemost.* **6**, 830–836 (2008).
21. Li, S. et al. F8 gene inversion and duplication cause no obvious hemophilia A phenotype. *Front Genet.* **14**, 1098795 (2023).
22. Chatron, N. et al. Severe Hemophilia A caused by an unbalanced chromosomal rearrangement identified using nanopore sequencing. *J. Thromb. Haemost.* **17**, 1097–1103 (2019).
23. Shi, M. et al. Clinical validation and application of targeted long-range polymerase chain reaction and long-read sequencing-based analysis for hemophilia: Experience from a hemophilia treatment center in China. *J. Thromb. Haemost.* **22**, 3431–3447 (2024).
24. Fahiminiya, S. et al. A full molecular picture of F8 intron 1 inversion created with optical genome mapping. *Haemophilia* **27**, e638–e640 (2021).
25. Fahiminiya, S. et al. Deciphering a novel complex inversion affecting F8 in a family with severe haemophilia A by optical genome mapping. *Haemophilia* **29**, 921–924 (2023).
26. Liu, B. et al. Optical genome mapping identified deletions, inversions, and insertions in hemophilia. *Blood Adv.* **9**, 360–364 (2025).
27. Lannoy, N., Lambert, C., Van Damme, A. & Hermans, C. Incidental finding of unreported large duplication in F8 gene during prenatal analysis: Which management for genetic counselling?. *Thromb. Res.* **182**, 39–42 (2019).
28. Jourdy, Y. et al. Complete characterisation of two new large Xq28 duplications involving F8 using whole genome sequencing in patients without haemophilia A. *Haemophilia* **28**, 117–124 (2022).
29. Lannoy, N. et al. Intron 22 homologous regions are implicated in exons 1–22 duplications of the F8 gene. *Eur. J. Hum. Genet.* **21**, 970–976 (2013).
30. Andersen, E. F., Baldwin, E. E., Ellingwood, S., Smith, R. & Lamb, A. N. Xq28 duplication overlapping the int22h-1/int22h-2 region and including RAB39B and CLIC2 in a family with intellectual and developmental disability. *Am. J. Med. Genet. A* **164a**, 1795–1801 (2014).
31. Ballout, R. A. et al. Int22h1/Int22h2-mediated Xq28 duplication syndrome: de novo duplications, prenatal diagnoses, and additional phenotypic features. *Hum Mutat.* **41**, 1238–1249 (2020).
32. Billes, A. et al. Possible incomplete penetrance of Xq28 int22h-1/int22h-2 duplication. *Clin Genet.* **106**(3), 234–246 (2024).
33. Zhang, S. et al. The GSA Family in 2025: A broadened sharing platform for multi-omics and multimodal data. *Genomics. Proteomics. Bioinform.* **23**(4), qzaf072 (2025).
34. CNCB-NGDC Members and Partners. Database resources of the national genomics data center, China national center for bioinformation in 2025. *Nucleic Acids Res.* **53**(D1), D30–D44 (2025).

Author contributions

Y.L. and Y.J. conceived and designed the study. Y.L., Y.J., X.Y., N.H., J.C., M.L. and K.Y. performed data curation. Y.L., Y.J., X.Y. and S.M. conducted formal analysis and investigation. Y.J., Y.L., Q.Q. and X.Z. provided resources. X.Y. created visualizations. X.Y. and Y.L. wrote the original draft. All authors reviewed the manuscript and approved the final version.

Funding

The study was supported by National Key Research and Development Program of China (2024YFC2707100), Peking Union Medical College Hospital Outstanding Young Talent Development Program C (UBJ10395), National High Level Hospital Clinical Research Funding (2022-PUMCH-B-076), and the National Key Clinical Specialty Construction Project (U114000).

Declarations

Competing interests

The authors declare no competing interests.

Consent to participate

Informed consent was obtained from all participants, and the study was approved by the Institutional Review Board of Peking Union Medical College Hospital (IRB no. I-25PJ0030). This study adhered to the principles of the Declaration of Helsinki.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-43654-x.

Correspondence and requests for materials should be addressed to Y.L. or Y.J.