**F8** genetic analysis strategies when standard approaches fail

B. Pezeshkpoor; A. Pavlova; J. Oldenburg; O. El-Maarri

Institute of Experimental Haematology and Transfusion Medicine, University of Bonn, Germany

**Keywords**
Haemophilia A, mutation negative patients, extended diagnostic flowchart

**Summary**
Haemophilia A is a common X-linked recessive disorder caused by mutations in **F8** leading to deficiency or dysfunction of coagulant factor VIII (FVIII). Despite tremendous improvements in mutation screening methods, in a small group of patients with FVIII deficiency suffering from haemophilia A, no DNA change can be found. In these patients, analysis reveals no causal mutations even after sequencing the whole coding region of **F8** including the flanking splice sites, as well as the promoter and the 3’ untranslated region (UTR). After excluding the mutations mimicking the haemophilia A phenotype in interacting partners of the FVIII protein affecting the half life and transport of the protein, mutations or rearrangements in non-coding regions of **F8** have to be considered responsible for the haemophilia A phenotype.

In this review, we present the experiences with molecular diagnosis of such cases and approaches to be applied for mutation negative patients.

**Correspondence to:**
Osman El-Maarri
Institute of Experimental Haematology and Transfusion Medicine, University of Bonn, Sigmund-Freud-Str. 25, 53127 Bonn, Germany
Tel. +49/(0)228/28716737, Fax +49/(0)228/28714320
E-mail: osman.elmaarri@ukb.uni-bonn.de

**Genanalytische Strategien für F8 bei Versagen der Standardverfahren**

Hämostaseologie 2014; 34: 167–173
http://dx.doi.org/10.5482/HAMO-13-08-0043
received: August 19, 2013
accepted in revised form: November 25, 2013
prepublished online: December 3, 2013

Haemophilia A is an X-linked bleeding disorder (OMIM 306700) with an occurrence of one in 5000 men (1). Typically, affected patients experience prolonged bleeding caused by lack or reduced residual activity of the coagulant factor VIII (FVIII:C). The severity of the disease is defined based on the amount of the residual FVIII:C level.

Haemophilia A is classified as
- severe (FVIII:C<1 IU/dl),
- moderate (1 IU/dl <FVIII:C<5 IU/dl), and
- mild (5 IU/dl<FVIII:C<40 IU/dl)
representing about 40%, 10%, and 50% of patients, respectively (2).

While most of the patients are referred to haemophilia centers based on a family history of haemophilia A, about one third of cases are sporadic. Haemophilia A in women is a rare observation (3–5). The female carriers are usually asymptomatic with bleeding events occurring only in about 10% of cases (6). The disease in carriers is mostly manifested when the carrier status is combined with skewed X chromosome inactivation (7, 8).

During the 1980s and early 1990s, mutation analysis of **F8** primarily focused on carrier detection (9, 10), while in the 1990s, genotype-phenotype associations became the driving force of **F8** analysis (11–14). Also, the impact of genetic factors on the risk of inhibitor development became more evident (15–17). To date, a wide range of mutations, scattered through the coding sequence of **F8**, have been reported that lead to either a qualitative or a quantitative defect of the FVIII protein.

The large and complex **F8** gene is located at the distal part of the long arm of the X chromosome (Xq28). It
- spans 186 kb of genomic DNA and
### Tab. 1 Diagnosis of haemophilia A: algorithm of tests

<table>
<thead>
<tr>
<th>Step</th>
<th>Assays and Tests for Haemophilia A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>FVIII:C&lt;sub&gt;act&lt;/sub&gt;/FVIII:C&lt;sub&gt;Clv&lt;/sub&gt;</td>
</tr>
<tr>
<td>2.</td>
<td>Inversion PCR, F8 sequencing, MLPA</td>
</tr>
<tr>
<td>3.</td>
<td>Long-range PCR, NGS, FISH, CGH, RT-PCR</td>
</tr>
<tr>
<td>4.</td>
<td>MLPA, FV activity measurement</td>
</tr>
<tr>
<td>5.</td>
<td>Genetic analysis of LMAN1 and MCFD2</td>
</tr>
<tr>
<td>6.</td>
<td>In silico analysis</td>
</tr>
</tbody>
</table>

#### Diagnosis and Mutation Detection Strategies

Generally, clinical diagnosis of haemophilia A is based on the FVIII:C level. Based on the residual FVIII:C, the severity of the disease is determined. This is critical for the choice of genetic analysis. The evaluation of the disease severity must be performed with caution; apart from preanalytical considerations (e.g., correct storage and amount of citrate anticoagulant, storage time) (24), the discrepancy between the results of one stage (FVIII:C<sub>act</sub>) and chromogenic (FVIII:C<sub>ch</sub>) assays have to be considered (25, 26). Thus, initial diagnosis of non-severe haemophilia A phenotypes should be based on results from both assays.

Most of the mutations in F8 (in exons and their flanking intron boundaries) can be identified using routine PCR-based methods with a mutation detection rate of up to 97% (23). The current mutation screening strategy applied for patients with a bleeding tendency and reduced FVIII:C is shown (Table 1).

1. In the first step, intron 22 (19, 20) and intron 1 (18) are screened for inversion mutations. This will identify the causal mutation in approximately 50% of severe haemophilia A patients.
2. In cases where no inversion is detected, the second step is the complete mutation analysis of F8 (exons and adjacent flanking intronic regions) to detect all mutations located in the exons of F8 as well as in their adjacent intronic regions.
3. The third step is the search for duplications and deletions using Multiplex Ligation-dependent Probe Amplification (MLPA). Duplications have been reported with an incidence of about 0.07% and can lead to mild-severe haemophilia A depending on their effect on the reading frame (27, 28).

**Mutation Negative Patients**

To elucidate the molecular mechanism behind FVIII deficiency in mutation negative patients further investigations beyond the above three steps are needed.

4. The forth step is to exclude misdiagnosis of haemophilia. Absence of mutations in genes that are known to be associated with FVIII deficiency provides further confirmation of the haemophilia A phenotype (Table 1).

5. In a fifth step, the integrity of the F8 locus has to be analyzed, as these regions might comprise complex rearrangements and/or DNA changes difficult to identify or escaping the current mutation screening protocols.

6. This last step is to clarify whether changes in intronic regions of F8 possibly affect mRNA splicing and transcription of mRNA.

As the last three steps are not routinely applied in genetic laboratories, the flowchart (Table 1) may simplify the choice for extended analysis to be applied in such cases.

---

**Note:**

- “consists of 26 exons ranging from 69 bp (exon 5) to 3106 bp (exon 14).
- The F8 coding sequence comprises only about 0.5% of the whole gene and encodes an mRNA of 9.1 kb.
- More than 2000 causal mutations in F8 are described in the haemophilia A mutation database (HAMSTeRS) (http://hadb.org.uk).

The molecular basis underlying haemophilia A is well characterized. In about half of severely affected patients, large genomic inversions are responsible for the disease. These inversions are the result of homologous recombination events between two copies of repeated DNA sequences in intron 1 (18) and intron 22 (19, 20) and their respective inverted homologous copies outside F8, and they have been reported in 1% (intron 1 inversion) and 40% (intron 22 inversion) of severe haemophilia A cases, respectively. One in 30,000 male infants are born with an inversion in F8 and a residual FVIII:C level below 1 IU/dl.

In the remaining cases, various types of point mutations (missense, nonsense and splice site mutations), small and large deletions, insertions, and duplications have been reported scattered throughout the F8 coding sequence, its flanking intronic regions and the 5’ and 3’ UTRs (21–23). However, in a small group of patients suffering from the “haemophilia A-like” phenotype, no detectable mutations in F8 can be found. To elucidate the molecular mechanism(s) leading to FVIII deficiency in these cases an extended molecular diagnostic procedure is needed.
Investigating the effect of known determinants of FVIII:C (step 4)

Causative mutations leading to haemophilia A are located within F8. However, mutations in interaction partners of the FVIII protein can lead to FVIII deficiency and mimic a „haemophilia A like“-phenotype. This includes mutations in
- lectin mannos-binding protein (LMAN1) (31) and
- multiple coagulation factor deficiency 2 protein (MCFD2) (32) as well as
- von Willebrand factor (VWF) (33–35).

Mutations in these genes could affect the secretion as well as the half-life of the FVIII protein (31, 36, 37). Genetic analysis of these genes avoids misdiagnosis of haemophilia A in patients suffering from combined FV/FVIII deficiency (38) and von Willebrand disease type 2 Normandy (VWD 2N).

To exclude combined FV/FVIII deficiency, determination of FV level is essential.

Combined FV/FVIII deficiency does not cause more bleeding than single deficiency of FV or FVIII, and patients suffer from a mild phenotype (31, 32, 36, 38). Generally, genetic analysis of LMAN1 and MCFD2 is applied when FV as well as FVIII levels are reduced. However, in patients with FVIII deficiency where no causal mutation in F8 is found, irrespective of normal FV activity levels, genetic analysis of LMAN1 and MCFD2 is still an appropriate target of investigation. Although no such cases have been described to date, the analysis could lead to the identification of novel mutations in LMAN1 and/or MCFD2 that may be specifically involved in the secretion of FVIII and not FV.

The next possible reason for FVIII deficiency is VWD type 2N caused by mutations in VWF. Such mutations result in normal VWF activity levels and a normal multimeric pattern, but reduced FVIII:C levels. VWF, a multimeric glycoprotein, is non-covalently bound to FVIII in plasma. Reduced binding of VWF to FVIII leads to FVIII deficiency caused by VWD 2N.

The disease is inherited as an autosomal recessive trait and is clinically similar to mild haemophilia A. The FVIII binding assay (37, 39) is the appropriate test for definitive diagnosis of VWD 2N in combination with sequencing the exons comprising the FVIII binding domains of VWF. For VWD 2N the minimum sequencing requirement should cover exons 18 to 20 However, isolated reports of mutations associated with a VWF:FVIII binding defect in exons 17 and 24–27 have been published (39).

Another approach in mutation negative patients, in addition to VWF:FVIII binding assay and genetic analysis of the FVIII binding domain of VWF, is the complete VWF mutation analysis. It might lead to the identification of new mutations causing weaker interaction of VWF and FVIII localized in other regions than the typical VWD 2N region of VWF.

Furthermore, several intra- and extragenic polymorphisms have been reported to modulate FVIII:C.
- A single polymorphism in exon 14 of F8 (rs1800291) has been reported to be associated with FVIII:C levels (40).
- Moreover, there is strong support for the association of AB0 blood type and FVIII:C (41, 42).
- Several genetic associations with FVIII:C and VWF:Ag levels have been reported using genome-wide association studies (GWAS) (43, 44).

However, the described SNPs influence FVIII:C levels to a lesser degree which would not result in an haemophilia A like phenotype. Even a combined occurrence of these SNPs is shown not to be associated with FVIII:C deficiency (45, 46).

Integrity of the F8 locus (step 5)

A wide spectrum of mutation mechanisms has been reported in patients classified as mutation negative patients. This includes both complex genomic rearrangements (47–49) as well as deep intronic mutations (46, 50). In this section, we describe possible gross new rearrangements of F8 and the ways to detect them.

Copy number variations (CNVs) and gross chromosomal rearrangements

Xq28 is prone, due to its high repetitive content (47, 51, 52), to frequent recombination events (45, 49, 50) (Tab. 1). Fluorescence in situ hybridization (FISH) is a widely used molecular cytogenetic technique in the diagnosis of diseases to locate a specific DNA sequence on the chromosomes. This approach can also be applied in haemophilia A diagnostics using a F8 specific probe. FISH analysis could exclude translocation and gross intra-chromosomal rearrangement of F8. However, only cases of haemophilia B in female patients which are due to translocation of F9 (53–55) have been described to date. Moreover, FISH analysis involves laborious cell culturing to generate metaphase chromosomes and can only detect rearrangements that involve >3 Mb of DNA (56).

Comparative genomic hybridization (CGH) (57) is a cost-effective alternative to the traditional cytogenetic methodology. CGH allows genome wide detection of duplication and deletions and helps to understand the role of genomic imbalance to a resolution of as low as 1kb (47, 57–59). It is a fast and robust method of genetic testing and has been applied for identification of deletions and duplications associated with haemophilia A (47, 52, 60). However, balanced chromosomal translocations and inversions with an occurrence of approximately 0.08% (61) are not detectable with the CGH array.

For diagnostic purposes, the described MLPA, due to its simplicity, is a good alternative for detection of CNVs within F8. In routine genetic analysis of haemophilia A, MLPA is mostly applied in female patients for detection of deletions and duplications and in male patients for identification of duplications (3, 27, 62).

Rearrangement involving F8 introns

The two recurrent mutation hotspots in intron 1 (18) and intron 22 (20) of F8 are prevalent examples of complex rearrangements in F8. Recently, a third inversion due to homologous recombination of inverted repeats within intron 1 of F8 was described in a single severe haemophilia A case (52). The three inversions occur due to the presence of homologous sequences in an in-
versions was developed by Rossetti (63). Recently, a new generation of strategies for the detection of intron 22 inversions evade standard PCR screening approaches. These above describe inversions in deep intronic regions of F8 coding for splicing errors has also been underlined in haemophilia A (50, 68, 69). The evidence of deeper intronic sequence variations affecting the splicing process is now becoming more obvious (65–67). The importance of screening for splicing errors has also been underlined in haemophilia A (50, 68, 69).

The most direct approach to identify DNA changes that might affect RNA splicing is qualitative reverse transcription-PCR (RT-PCR) (Tab. 1). The effect of several intronic variants in F8 located close to or within the splice site consensus sequences has been analyzed by several groups (69, 70). These groups have shown that not only changes close to or within the consensus splice sites but also single nucleotide substitutions in deep intronic regions of F8 may lead to alternative splicing. Such aberrant mRNAs mostly lead to haemophilia A due to the generation of premature stop codons (Tab. 2) (46, 50, 71). Such transcripts would lead to a production of a truncated protein or coupled surveillance degradation of the mRNA (72).

The evidence of the pathogenic impact of such deep intronic mutations in haemophilia A was first reported in 1999 when Bagnall et al. described a mutation in intron 1 of F8 (c.143+1567A>G) that activates a cryptic exon and generates a premature stop codon in F8 mRNA (73). This mutation was found in two patients suffering from severe and moderate haemophilia A phenotype, with FVIII:C of 1 IU/dl and 2 IU/dl, respectively. Other deep intronic mutations also lead to aberrant mRNA splicing. However, RT-PCR analysis for this mutation shows the availability of the wild type F8 transcript along with the aberrant transcript, which leads to a milder haemophilia A phenotype (Tab. 2).

Sequencing the whole F8 locus is a possible approach for further investigations in patients where no alternative splicing of mRNA is observed. For this purpose, a PCR fragment can be obtained by LR-PCR approach and sequenced using a next generation sequencing (NGS) platform (Tab. 1). The advantage of NGS is that the whole F8 locus is screened for variations in one single step. This approach might identify intronic variants which may fail to show an effect on RNA splicing due to several limitations of the currently used RT-PCR approach (46) (e. g. the SNP might overlap with the primers used for the reverse transcription or the splicing effect would evade the nested RT-PCR).

However, to identify patient specific variants, all intronic variations need to be filtered using the SNP database (dbSNP) and the 1000 genomes project (www.ncbi.nlm.nih.gov/SNP/ and www.1000genomes.org/). Furthermore, a panel of healthy individuals has to be screened to exclude rare polymorphisms related to ethnicity of the patient cohort. All variants that pass these exclusion criteria and are recurrent in this group of patients may be considered as potential intronic mutations. However, the causality of such mutations

| Tab. 2 Deep intronic mutations causing mild haemophilia A in mutation-negative patients |
| mutation in F8 | intron position | experimental consequence: inclusion of | effect | reference | patients (n) |
| c.143+1567A>G | 1 | 191 bp | pre mature stop codon | (73) | 2 |
| c.1537+325A>G | 10 | 226 bp | | (71) | 1 |
| c.2113+1152delA | 13 | 112 bp | | (50) | 1 |
| c.5587–93C>T | 16 | 56 bp | | (46, 50) | 1 |
| c.5998+530C>T | 18 | 105 bp | | (46) | 1 |
| c.5998+941C>T | 126 bp | | | (46) | 1 |
| c.5999–227G>A | 35 bp / 55 bp | in-frame protein with extra 30 aa | | (50) | 3 |

aa: amino acids
needs to be proven by functional assays (e.g. luciferase assay).

While the RT-PCR approach is the initial method to screen for abnormal splicing, the NGS approach has the advantage of screening all introns in one step thus localizing all potential changes to be investigated in detail using specific RT-PCR approaches.

Patients without apparent causative mutation

And yet, in some patients, a causal mutation explaining the haemophilia A phenotype cannot be detected. In this group, the reason for FVIII deficiency remains elusive. Several possibilities must be investigated to identify the molecular mechanisms behind the disease in these cases.

The first one is tissue mosaicism.

A point mutation in F8 could be present only in the liver among other tissues but not in blood. Although mutations causing haemophilia A usually appear to have arisen in germ cells, a de novo mutation may also occur during early embryogenesis and thus may represent either germine and/or somatic mosaicism (74–78). However, this is only true in sporadic cases, but not for patients with a family history of haemophilia A.

The second possibility is rapid degradation of FVIII protein or its rapid uptake in blood.

Previous studies have shown that LDL receptor-related protein (LRP-1) and other members of the LDL receptor family affect the clearance of FVIII (79). Moreover, C-type lectin domain family 4 member M (CLEC4M), a calcium-dependent mannose-specific receptor, was recently shown to influence clearance of VWF (80). CLEC4M is expressed in the main site of FVIII production, the liver sinusoidal endothelial cells, and due to the important role of VWF in protection of FVIII from degradation, such mutations or polymorphisms may indirectly contribute to variability in FVIII:C. While it is possible that mutations at other loci play a role in intracellular trafficking and/or secretion of FVIII molecule, it remains unclear whether such mutations lead to FVIII deficiency.

Another possibility is a putative regulatory region, the so-called locus control region (LCR), located within or upstream of F8.

Recent studies on several loci, i.e. β-globin, have offered new insights in support of long distance interaction between enhancers and promoters from 6 to 25 kb upstream of the gene (81). Such LCRs are defined functionally due to their ability to influence the expression of linked genes. The fact that the defect leading to haemophilia A like phenotype in most cases indicates a X-chromosomal inheritance pattern makes the role of such LCRs in regulation of F8 expression more convincing.

In order to elucidate the exact molecular mechanism behind haemophilia A phenotype in the remaining patients, more genetic and cellular approaches need to be applied. For example,

- a genome wide association study may be carried out if several affected members from the same family are available,
- a cellular approach by establishing a patient specific cellular system.

The improvements in modeling the diseases by induced pluripotent stem cells could facilitate a cellular system to understand and investigate the molecular basis of haemophilia.

Conclusion

Despite remarkable improvements in mutation detection technologies, in a small group of patients with reduced FVIII:C, the cause of haemophilia cannot be found in the exons of F8.

In such cases, it is necessary to look beyond the coding sequences, exon/intron boundary, and promoter region.

Here, a systematic screening protocol is needed to identify the molecular cause of the disease.

In this review, we summarize our experience in analyzing such cases and we describe a six steps protocol to identify causal mutations. We present several techniques for successful identification and confirmation of novel mutations. These techniques may be applied not only in haemophilia A patients, but also in the genetic analysis of other monogenic diseases, where no pathogenic mutation is found in the coding sequences.

Acknowledgment

This work was supported by a DFG grant (EL499/2–1), a Baxter bioscience grant (H12–000820) and a Bayer Haemophilia Awards Program.

Conflict of interest

The authors declare that they have no conflict of interest.

References