Deep intronic variants in the factor VIII gene

J Elisa Bach (1), Beat Wolf (2,3), Johannes Oldenburg (4), Clemens R Müller (1), Simone Rost (1)

(1) Department of Human Genetics, University of Würzburg, Germany
(2) University of Applied Sciences Western Switzerland, Switzerland
(3) Department of Bioinformatics, University of Würzburg, Germany
(4) Institute of Experimental Haematology and Transfusion Medicine, University of Bonn, Germany

INTRODUCTION

Haemophilia A (HA) is the most common X-linked bleeding disorder with an incidence of 1 in 5000 males. Different mutation types in the factor VIII gene (*F8*) lead to a decrease in coagulation factor VIII activity and variable severities of the patients' phenotypes. The causative alteration can be detected in the vast majority of HA patients by standard diagnostic screening methods targeting the coding regions of the *F8* gene. Still, this approach fails in about 4 % of cases. This study intended to analyse the whole *F8* gene including all intronic sequences in 16 HA patients by next generation sequencing (NGS) in order to screen for deep intronic variants.

METHODS

Fifteen of the male index patients showed mild and one case a moderate HA phenotype with no abnormality in the coding sequences and splice sites of the *F8* gene identified by standard diagnostic techniques. Patient 16 was pre-diagnosed with a duplication of exons 1-22 and a triplication of exons 23-25 of the *F8* gene by multiplex ligation-dependent probe amplification (MLPA). Target enrichment for the whole genomic region of *F8* was performed with the SureSelectXT system (Agilent) and sequencing on a MiSeq (Illumina). NGS data were analysed with GenSearchNGS (PhenoSystems).

RESULTS

NGS data revealed 23 deep intronic candidate variants in different *F8* introns. Three of the single nucleotide variants (SNVs) have been described before as HA causing mutations. Six other variants were recurrent among the 15 cases studied. Several bioinformatic tools were used to score all candidate variants regarding their potential deleteriousness (e.g. by C-scores, Fig. 1) and the creation of de novo splice sites (Fig. 2), also in comparison to already published deep intronic *F8* mutations. In each of the patients, at least one SNV or copy number variation (CNV) was predicted to be potentially pathogenic.

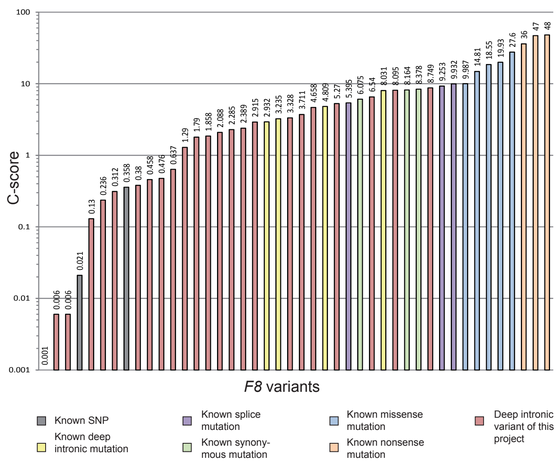

Figure 1: C-scores for the 23 variants found in this study as compared to several known variant types in the *F8* gene (SNPs, deep intronic mutations, splice mutations, synonymous mutations, missense and nonsense mutations causing HA). Scores were calculated with the Combined Annotation Dependent Depletion (CADD) analysis predicting the deleteriousness of SNVs by combining different functional aspects and diverse annotations into a single parameter, the C-score).

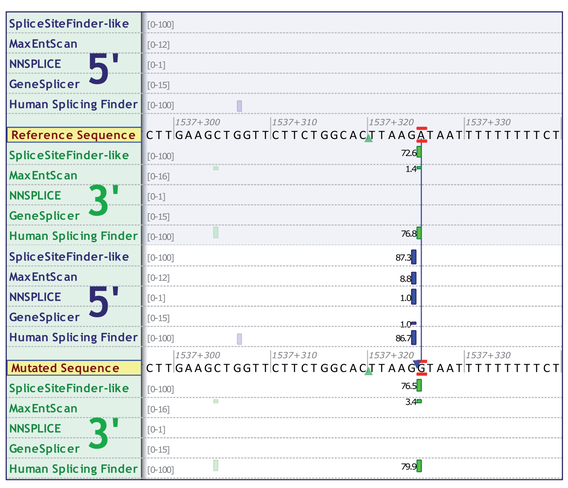

Figure 2: Splice prediction by Alamut software (Interactive Biosoftware) for the wildtype sequence on the top (highlighted grey) and mutated sequence with c.1537+325A>G in intron 10 beneath (white). Five splice prediction tools indicate a high probability for the use of the mutated site as *de novo* donor splice site. The mutation had already been published as pathogenic by Inaba *et al.* (2013) and was found in one patient of our study.

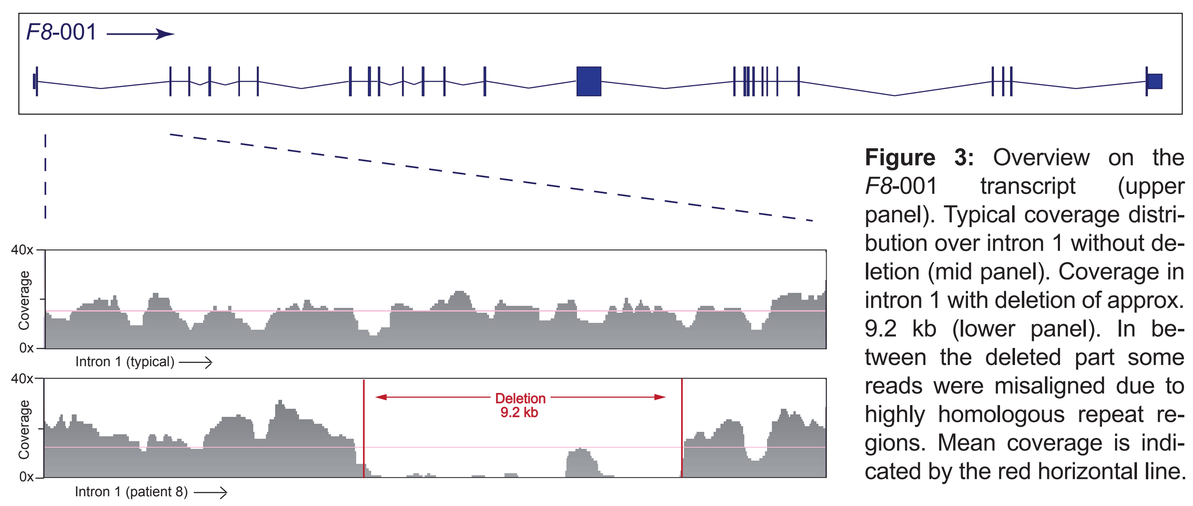

Figure 3: Overview on the *F8*-001 transcript (upper panel). Typical coverage distribution over intron 1 without deletion (mid panel). Coverage in intron 1 with deletion of approx. 9.2 kb (lower panel). In between the deleted part some reads were misaligned due to highly homologous repeat regions. Mean coverage is indicated by the red horizontal line.

The patient with moderate HA additionally showed a deletion of 9.2 kb in intron 1 with both breakpoints located in Alu-type repeats (Fig. 3). The CNVs of patient 16 could be well seen in the NGS data by comparing the coverage of the affected exons, but it wasn't possible to define exact breakpoints (Fig. 4).

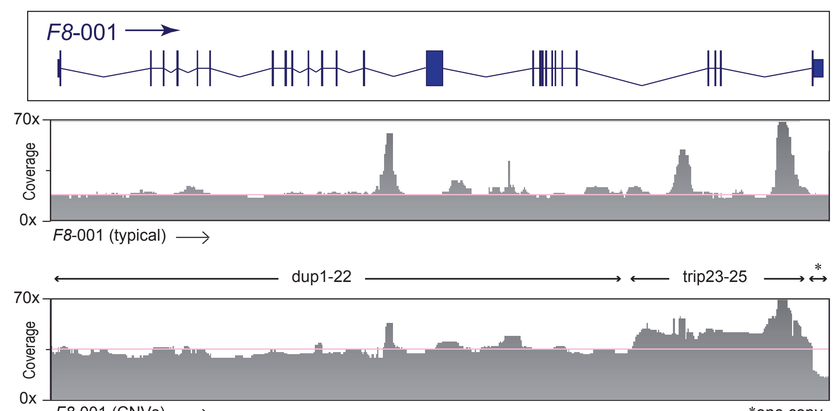

Figure 4: Overview on the *F8*-001 transcript (upper panel). Typical coverage distribution over the whole *F8* gene with normal dose, i.e. one copy per exon (mid panel). Copy number variations of this patient with duplication of exons 1-22, triplication of exons 23-25 and only exon 26 with normal dose (lower panel). Mean coverage is indicated by the red horizontal line.

CONCLUSION

In summary, this NGS approach proved an effective method to analyse the whole *F8* gene for potentially pathogenic deep intronic variants and CNVs in a selection of 16 male HA patients. In general, this approach bears the potential to be applied as efficient one-step sequencing method of the complete *F8* gene in molecular diagnostics of haemophilia A. Yet, in a diagnostic setting besides in silico prediction further functional studies like mRNA analysis would be required to confirm the causality of deep intronic variants. As this study was pseudonymized, mRNA analyses on patients' blood samples were not possible. Therefore, in a recent study we successfully tested mini-gene assays to confirm causality of some of the candidate variants.

References

1. JE Bach, B Wolf, J Oldenburg, CR Müller, S Rost: *Identification of deep intronic variants in 15 haemophilia A patients by Next Generation Sequencing of the whole factor VIII gene.* Thromb Haemost (2015)
2.. JE Bach, CR Müller, S Rost: *Mini-gene assays confirm the splicing effect of deep intronic variants in the factor VIII gene.* Thromb Haemost (Submitted)
3. JE Bach, J Oldenburg, CR Müller, S Rost: *Mutational spectrum and identification of deep intronic variants by next generation sequencing in the factor VIII gene of haemophilia A patients.* Haemostaseologie (Submitted)
4. Inaba H, Koyama T, Shinozawa K, et al. *Identification and characterization of an adenine to guanine transition within intron 10 of the factor VIII gene as a causative mutation in a patient with mild haemophilia A.* Haemophilia (2013)

contact: elisa.bach@uni-wuerzburg.de