Flipping the tip of the X

Edward G. D. Tuddenham

Haemostasis Research Group, Clinical Research Centre, Harrow HA1 3UJ, UK

The commonest severe bleeding disorder in man is haemophilia A — a classic X-linked disorder. Although a male specific bleeding syndrome has been known for at least two millennia, Factor VIII (FVIII), the protein whose deficiency causes the bleeding, was only purified to homogeneity in 1983[1]. A co-factor in the intrinsic pathway of blood coagulation, it is present in blood in only minute amounts, is unstable and is also a minor component of the carrier protein complex in which it circulates — FVIII/von Willebrand factor.

Purification of the protein led rapidly, thanks to the involvement of the biotechnology industry, to the cloning of the FVIII gene, *in vitro* expression of functional FVIII[2–5] and not long after, identification of FVIII mutations[6]. Finding such mutations was however a hit and miss affair, due to the size and complexity of the gene which, with 26 exons spanning 189 kilobases of genomic DNA (in Xq 2.8) was at the time the largest reported gene.

In the absence of methods for rapidly scanning DNA for point mutations detection of mutations proceeded slowly in the mid 80s. Those that were reported mostly affected *Taq*I sites or were large deletions detected by Southern blotting. These two types of mutations occur in about 5–10% of randomly screened patients with haemophilia A. Being an X-linked lethal or sub-lethal disorder, haemophilia A must be maintained in the population by new mutations and the mutations are expected to be diverse. The advent of rapid PCR based screening methods confirmed this prediction and led to an increase in the number of mutations detected[7]. Thus Higuchi *et al.*[8] amplified each of the 26 exons and flanking regions of the factor VIII gene by PCR and screened them by denaturing gradient gel electrophoresis. A surprising result of this study was that whereas most patients with moderate or mild disease had missense mutations, and many severe cases had nonsense or frameshift mutations, no mutation at all could be detected in half of the severely affected patients. This mystifying result prompted several hypotheses but an explanation only began to emerge when Naylor and colleagues[9] screened reverse transcribed ectopic messenger RNA for factor VIII. They found that in half of their severely affected patients, there was no abnormality in the exons. However, a defect in the splicing of exon 22 to exon 23, was demonstrable in these cases. Between these two exons lies the largest (32.1 kb) intron of the factor VIII gene. This intron is of further interest, since it contains a CpG island, from either side of which two additional transcripts originate. The first of these, F8A, is a wholly contained intronless gene, which it later transpired, also occurs in at least two copies near the telomere of the X-chromosome[10,11]. The second transcript, F8B, encodes eight additional amino acids and uses exons 23–26 of the FVIII gene[12]. Although both these transcripts are found in all tissues, they are of unknown function.

With this information Gitschier and colleagues[13] reasoned that a possible explanation for Naylor's observation could be as follows. An inversion involving homologous recombination between the extragenic and intragenic copies of the F8A, would lead to a separation of the two halves of the gene. However transcripts corresponding to exons 1–22 and 23–26 would still be found due to the presence of the endogenous factor VIII promoter for the former and F8B for the latter. In an exciting paper in this issue of *Nature Genetics*[13], this hypothesis is convincingly proved by satisfyingly straightforward methods. A probe corresponding to part of intron 22 in the region of F8A is used to probe a Southern blot of genomic DNA digested with *Bcl*I. Rearrangements attributable to an inversion involving either extragenic copy of F8A were found. Furthermore, the rearrangement was shown to occur in nearly half of all cases of severe haemophilia A. Next they identified novel sequences attached to the 5' *FVIII* transcript of such a case. Then using YAC clones they located the novel sequences close to the telomeric F8A copies, thus confirming the hypothesis since a partial *FVIII* transcript could only be spliced to such sequences by a major inversion involving intron 22. Naylor and colleagues[14] have independently reached the same conclusion using a similar approach.

The practical effect of these studies is that clinical genetics departments confronted with a kindred containing a sporadic case of haemophilia A can now confidently set about locating a mutation. Mutation specific diagnosis is especially useful where a colateral or antecedent female relative of a sporadic case requires genetic counselling since, by definition the level of the mutation is unknown. Therefore with this new information an optimum strategy would be to "invert" the usual order and screen at once for the inversion mutation in all severe cases (factor VIII less than 1%) using the F8A probe against *Bcl*I digested DNA. In the absence of a rearrangement, exon screening, and then as a last resort (family structure permitting) linkage analysis using the highly polymorphic CA repeat within intron 13(ref. 15) could be used.

This type of inversion appears so far to be unique to factor VIII, no doubt reflecting the curious distribution of the F8A gene with both intronic and extragenic copies. If tissue were available from Queen Victoria or one of her affected descendants, the chances are about 40% that the flip tip inversion could be detected, settling the medico-historical mystery — was it factor VIII or factor IX deficiency that afflicted the Royal families of Europe? □

References

1. Rotblat, F. *et al.* Thrombosis & Haemostasis **50**, 108 (1983).
2. Gitschier, J. et al. *Nature* **312**, 326–330 (1984).
3. Wood, WI. *et al. Nature* **312**, 330–337 (1984).
4. Vehar, GA. *et al. Nature* **312**, 337–342 (1984).
5. Toole, JJ. *et al. Nature* **312**, 342–347 (1984).
6. Gitschier, J. *et al. Nature* **315**, 427–430 (1985).
7. Tuddenham, E. *et al. Nucl. Acids Res.* **19**, 4821–4833 (1991).
8. Higuchi, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 7405-7409 (1991).
9. Naylor, JA. *et al. Hum. molec. Genet.* **2**, 11–17 (1993).
10. Levinson, B. *et al. Genomics* **13**, 862–865 (1992).
11. Freije, D & Schlessinger, D. *Am. J. hum. Genet.* **51**, 66–80 (1992).
12. Levinson, B. *et al. Genomics* **14**, 585–589 (1992).
13. Lakich, D. *et al. Nature Genet.* **5**, 236–241(1993).
14. Naylor, J. *et al. Human molec. Genet.* (in the press).
15. Lalloz, MR. *et al. Lancet* **338**, 207–211 (1991).