REVIEW

Haemophilia A and haemophilia B: molecular insights

D J Bowen

J Clin Pathol: Mol Pathol 2002;**55**:1–18

This review focuses on selected areas that should interest both the scientist and the clinician alike: polymorphisms within the factor VIII and factor IX genes, their linkage, and their ethnic variation; a general assessment of mutations within both genes and a detailed inspection of the molecular pathology of certain mutations to illustrate the diverse cause–effect relations that exist; a summary of current knowledge on molecular aspects of inhibitor production; and an introduction to the new areas of factor VIII and factor IX catabolism. An appendix defining various terms encountered in the molecular genetics of the haemophilias is included, together with an appendix providing accession numbers and locus identification links for accessing gene and sequence information in the international nucleic acid databases.

The molecular basis of the hereditary haemophilias has been unravelled to a remarkable degree over the past decade. The early studies, which glimpsed the possibility of mutational heterogeneity, have been verified by the almost breath taking diversity of mutations that have since been characterised in the coagulation factor VIII gene in haemophilia A and the coagulation factor IX gene in haemophilia B. Although mutational heterogeneity abounds within these two genes, natural polymorphic heterogeneity does not. These apposite features have made the collection of informative genetic data for familial investigations a difficult task. However, modern techniques of molecular biology have decreased the enormity of the task of mutation analysis in these two genes and have decreased the workload for the analysis of their polymorphisms. This review focuses on polymorphisms and mutations in the genes for coagulation factor VIII and factor IX, molecular biological aspects of inhibitor production, and the new areas of factor VIII and factor IX catabolism. The molecular pathology of certain mutations is dealt with to exemplify the interesting and revealing cause–effect relations that exist within the haemophilias.

> "An almost breath taking diversity of mutations have since been characterised in the coagulation factor VIII gene in haemophilia A and the coagulation factor IX gene in haemophilia B"

Correspondence to:
D J Bowen, Department of Haematology, University of Wales College of Medicine, Heath Park, Cardiff CF3 0DQ, South Wales, UK;
bowendj1@cf.ac.uk

Accepted for publication
10 May 2001

BACKGROUND: FACTOR VIII AND FACTOR IX IN COAGULATION

These two proteins circulate as inactive precursors that are activated at the time of haemostatic challenge, via the intrinsic or extrinsic pathways of the coagulation cascade (fig 1). Factor VIII is a protein cofactor with no enzyme activity per se; factor IX is a serine protease with an absolute requirement for factor VIII as cofactor. Upon activation, and in the presence of calcium ions and phospholipid surfaces, factor VIII and factor IX form an active complex, the tenase complex, which activates factor X (fig 1). Subsequent stages of the cascade then proceed, culminating in the deposition of fibrin, the structural polymer of the blood clot.

A deficiency or dysfunction of either factor VIII or factor IX compromises the activation of factor X, so that the ensuing steps of the coagulation cascade are also compromised and fibrin deposition is either inefficient or non-existent. Thus, the fundamental biochemical lesion underlying the haemophilias is an insufficiency of the activity of the tenase complex, brought about either by a deficiency of coagulation factor VIII cofactor activity (haemophilia A) or coagulation factor IX enzyme activity (haemophilia B). Thus, it is not surprising that the two disorders are clinically similar because they both arise from perturbation of the same essential step in the process of fibrin generation.

THE FACTOR VIII AND FACTOR IX GENES AND PROTEINS

Both genes are located on the X chromosome: the factor VIII gene is located towards the end of the long arm at Xq28[1]; the factor IX gene is also located on the long arm, more towards the centromere at Xq27.[2] There is a considerable genetic distance (0.5 cM) between the two genes[3 4]; they are by no means in tight linkage with each other. Thus, haemophilia A and B are X linked disorders; they are recessive, carried by females (karyotype 46:XX), and present in males (karyotype 46:XY). Rare cases of female haemophilia are known[5 6] and these may arise through the presence of two defective factor VIII or factor IX genes, or through non-random X chromosome inactivation. The latter has often been offered as a primary explanation for haemophilia symptoms in known carriers; however, recent data suggest that there is an absence of correlation between X chromosome inactivation and plasma concentrations of factor VIII in carriers of haemophilia A and B.[7]

Abbreviations: ARE, androgen response element; C/EBP, CCAAT/enhancer binding protein; EGF, epidermal growth factor; LD, long distance; LRP, lipoprotein receptor related protein; MHC, major histocompatibility complex; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; SNP, single nucleotide polymorphism; VNTR, variable number tandem repeat sequence

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

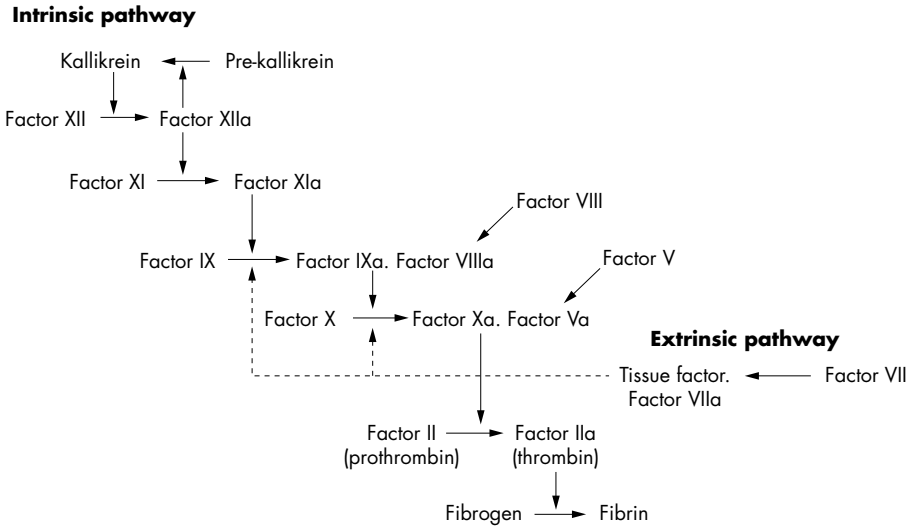


Figure 1 Schematic showing the intrinsic and extrinsic pathways of the coagulation cascade leading to fibrin formation. A deficiency or dysfunction of coagulation factor VIII or factor IX compromises the activation of factor X, the ensuing reactions are inefficient and haemophilia results.

Factor VIII

The factor VIII gene is extremely large (~180 kb) and structurally complex (26 exons) (fig 2A; table 1).[1] Intron 22 of the gene contains a CpG island, which acts as a bidirectional promoter for two further genes, F8A[8] and F8B[9] (fig 2B). The CpG island and F8A are contained within a stretch of DNA of approximately 9.5 kb, which is repeated at least twice on the X chromosome, further towards the telomere and extragenic to the factor VIII gene[8 10] (fig 2C). These homologues are known as int22h-1 (intragenic) and int22h-2 and int22h-3 (extragenic).[10]

F8A is intronless and spans less than 2 kb within intron 22.[8] It is transcribed in the direction opposite to factor VIII.[8] It is transcribed abundantly in a wide variety of cell types[8] and is conserved in the mouse, which implies it has some function.[11]

F8B comprises a 5′ exon located within intron 22 and then exons 23–26 of the factor VIII gene.[10] The 5′ exon potentially encodes eight amino acids and the factor VIII reading frame is maintained for exons 23–26.[10] The result is a protein that is considerably shorter than factor VIII and that includes the phospholipid binding domain.

The factor VIII gene transcript is approximately 9010 bases in length and comprises a short 5′ untranslated region (150 bases), an open reading frame plus stop codon (7056 bases), and an unusually long 3′ untranslated region (1806 bases)[12] (fig 3A). The open reading frame encodes a signal peptide of 19 amino acids, which directs the passage of factor VIII through the cell, and a mature protein of 2332 amino acids, which is the circulating inactive pro-cofactor (fig 3B).[12] This contains a distinct domain structure based on homologies within the protein.[13]

Factor VIII is activated by proteolysis catalysed by thrombin (fig 3C) (reviewed in Tuddenham and Cooper[14]). The principal activation cleavage sites are on the C-terminal side of arginine residues 372, 740, and 1689.[15] Activated factor X also cleaves at these sites. The activation cleavage sites flank the B domain, which is released from factor VIII on activation, leaving a heterotrimer comprising an N-terminal heavy chain and a

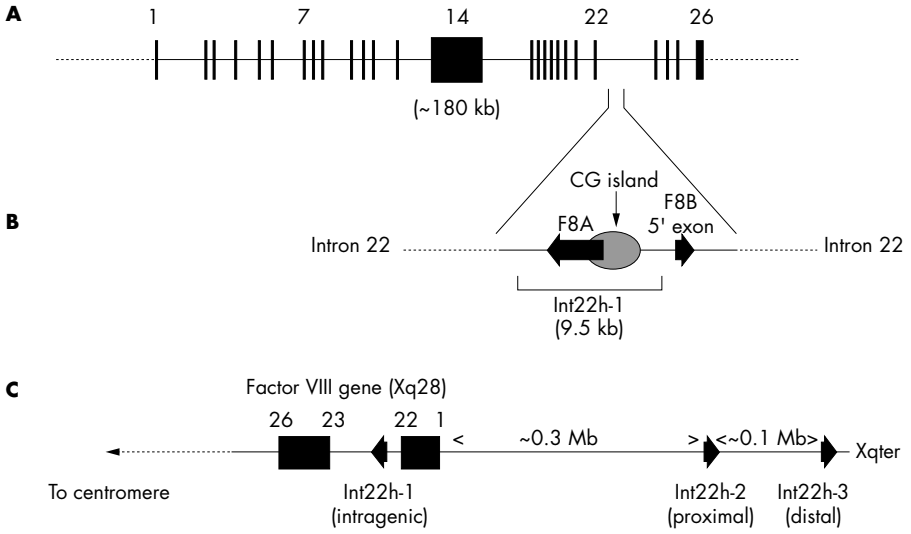


Figure 2 (A) Genomic organisation of the human factor VIII gene. Exon and intron sizes are given in table 1. (B) Enlargement of a portion of intron 22 of the factor VIII gene to show the relative locations and orientations of the CG island, the intragenic gene F8A, the 5′ exon of the putative gene F8B, and the extent of the int22h-1 homologue, which is duplicated towards the telomere of Xq. (C) Orientation of the factor VIII gene within Xq28 and location and orientation of int22h homologues.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

Table 1 Exon and intron sizes for the human factor VIII gene[1]

Exon	Length (bp)	Intron	Length (kb)
1	313	1	22.9
2	122	2	2.6
3	123	3	3.9
4	213	4	5.4
5	69	5	2.4
6	117	6	14.2
7	222	7	2.6
8	262	8	0.3
9	172	9	4.8
10	94	10	3.8
11	215	11	2.8
12	151	12	6.3
13	210	13	16.0
14	3106	14	22.7
15	154	15	1.3
16	213	16	0.3
17	229	17	0.2
18	183	18	1.8
19	117	19	0.6
20	72	20	1.6
21	86	21	3.4
22	156	22	32.4
23	145	23	1.4
24	149	24	1.0
25	177	25	22.4
26	1958	26	

C-terminal light chain; these are held together by Ca^{2+}. In the circulation, factor VIII is carried and protected by von Willebrand factor.[16 17] Proteolytic cleavage of factor VIII at the time of activation simultaneously releases it from its complex with von Willebrand factor.[18 19]

"In the circulation, factor VIII is carried and protected by von Willebrand factor"

Factor IX

The factor IX gene is considerably smaller and structurally simpler than that for factor VIII. It is approximately 34 kb in length and contains only eight exons, the largest of which is only 1935 bp (fig 4A; table 2).[2] The transcript is 2803 bases in length and comprises a short 5′ untranslated region (29 bases), an open reading frame plus stop codon (1383 bases), and a 3′ untranslated region (1390 bases) (fig 4B).[20] The open reading frame encodes a pre-pro-protein in which the pre-sequence (or signal sequence) directs factor IX for secretion, the pro-sequence provides a binding domain for a vitamin K dependent carboxylase, which carboxylates certain glutamic acid residues in the adjacent GLA domain, and the remainder represents the factor IX zymogen, which enters the circulation after removal of the pre- and pro-sequences. Domains within the zymogen are identified according to structure or function.

Activation of factor IX involves cleavage of two peptide bonds, one on the C-terminal side of arginine 145 (the α-cleavage) the other on the C-terminal side of arginine 180 (the β-cleavage) (fig 4C).[21] These cleavages are brought about by activated factor XI generated through the intrinsic pathway,[22] or via the tissue factor/activated factor VII complex of the extrinsic pathway.[23 24] The relative importance of these two pathways to factor IX activation is indicated by the fact that severe factor XI deficiency is typically associated with a mild bleeding disorder, whereas severe factor VII deficiency is characterised by a severe bleeding disorder. The activation cleavages generate an N-terminal light chain and a C-terminal heavy chain, held together by a disulphide bond between cysteine residues 132 and 279 (fig 4D).[21]

The above overview is necessarily brief. Its purpose is to relate salient features that provide the context and framework for the molecular information that follows. More in depth coverage can be obtained from several excellent reviews.[14 25–28]

POLYMORPHISMS

The factor VIII and factor IX genes both contain two types of polymorphism: single nucleotide polymorphisms (SNPs) and length polymorphisms, also known as variable number

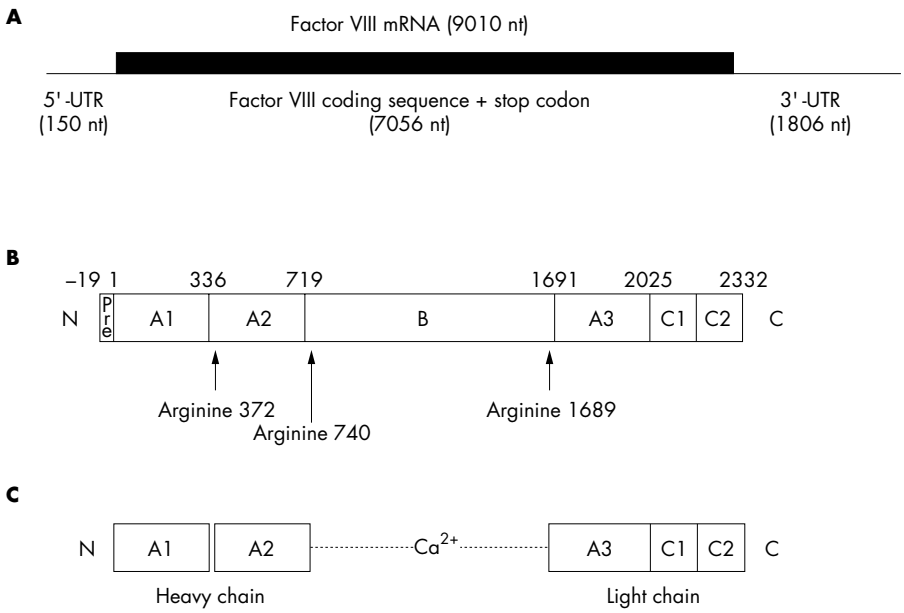


Figure 3 (A) Factor VIII mRNA showing the extent and location of the open reading frame. (B) The newly synthesised factor VIII protein molecule comprising a pre-sequence of 19 amino acids and a mature peptide of 2332 amino acids (total length, 2351 amino acids). A1–3, B, C1, and C2 represent domains assigned according to homologies within the protein (the boundaries are somewhat arbitrary). The arginine residues signalling the sites for proteolytic activation are arrowed. (C) Activated factor VIII comprising a heterotrimer in which the dimeric N-terminal heavy chain is held together with the monomeric C-terminal light chain by a metal ion bridge (Ca^{2+}).

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

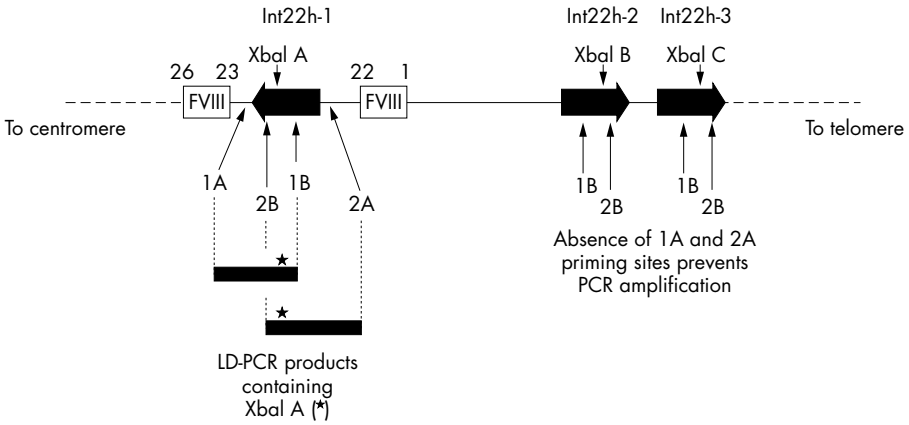


Figure 4 (A) Genomic organisation of the human factor IX gene. Exon and intron sizes are given in table 2. (B) Factor IX mRNA showing the relative size and location of the open reading frame. (C) The newly synthesised factor IX protein molecule comprising a pre- and pro-sequence (27 and 19 amino acids, respectively) and a mature peptide of 415 amino acids (total length, 461 amino acids). (D) Activated factor IX comprising an N-terminal light chain and a C-terminal heavy chain held together by a disulphide bridge between cysteine resides 132 and 279. GLA, "GLA" domain, in which 12 glutamic acid residues undergo post-translational γ-carboxylation by a vitamin K dependent carboxylase; EGF, epidermal growth factor-like domain; act, activation peptide released after proteolytic activation at arginine 145 and arginine 180; catalytic, the serine protease domain.

tandem repeat sequences (VNTRs) or microsatellites. Polymorphisms have a scientific interest of their own and are useful in areas as diverse as forensic science[29 30] and the study of human evolution.[31 32] They have clinical relevance in the context of hereditary disorders in that they can be used to track a defective (or normal) gene through an affected family. Such linkage studies have permitted carrier status investigation and prenatal diagnosis in haemophilia A and haemophilia B.[33 34]

There is an apparent paucity of polymorphisms in both the factor VIII and factor IX genes. In the recent presentation of the sequence of the human genome,[35] two haploid human genomes chosen randomly differed by 1 base/1250 bases on average; therefore, the factor VIII and factor IX genes would be expected to contain approximately 144 and 27 SNPs, respectively. However, these numbers are considerably greater than the number of polymorphisms detected so far in these genes, although this could reflect a problem of detection rather than a genuine paucity. Certainly, many more candidate polymorphisms have been identified as a result of the human genome sequencing programme[35 36] than the recognised polymorphisms described below.

"There is an apparent paucity of polymorphisms in both the factor VIII and factor IX genes"

FACTOR VIII GENE POLYMORPHISMS

The factor VIII gene contains several SNPs, many of which fall into the subcategory of restriction fragment length polymorphisms (RFLPs) (fig 5A). In addition, two microsatellites have been reported within the gene, one in intron 13,[38] the other in intron 22 (fig 5A).[44]

Table 2 Exon and intron sizes for the human factor IX gene[2]

Exon	Length (bp)	Intron	Length (bp)
1	117	1	6206
2	164	2	188
3	25	3	3689
4	114	4	7163
5	129	5	2565
6	203	6	9473
7	115	7	668
8	1935		

The intron 13 microsatellite comprises the simple repeat (CA)n,[38] whereas the intron 22 comprises (CA)n(CT)n.[44] Differences in the value of "n" at such repeats have been attributed to DNA polymerase slippage and appear to drift towards an increase in repeat length.[57] Because such events could occur between generations within a family, microsatellite data should be interpreted with appropriate caution in a linkage study.[58] Preferably, verification of microsatellite data should be sought from at least one other intragenic locus.

The XbaI and MspI RFLPs in intron 22 of the factor VIII gene (designated XbaI A and MspI A) occur within the int22h-1 sequence.[10 42 43] The corresponding positions in the extragenic homologues int22h-2/3 are also polymorphic for XbaI[42] and MspI[43] (designated XbaI B and C, and MspI B and C, respectively). Caution must therefore be exercised in the use of these two highly informative loci.

The allele frequencies of some of the polymorphisms have been investigated in various populations and differences are seen (table 3). Three observations are of particular interest: first, the [+] alleles for the BclI and HindIII RFLPs have very different frequencies in American blacks compared with other ethnic groups, affording a small glimpse into human evolutionary divergence. Second, although whites and Chinese have similar allele frequencies at the BclI, HindIII, and XbaI A loci, they differ noticeably at the BglI locus. The BglI RFLP is the result of a transition that is not likely to recur independently.[62] Thus, it is possible that its presence in whites but not in Chinese might reflect its emergence in the former after the two ethnic groups diverged. However, the presence of the BglI RFLP in Japanese and American blacks suggests that this explanation is too simplistic. Third, in American blacks the allele frequencies for the BclI and HindIII loci are the reverse of what is seen in other ethnic groups; however, for the BglI locus, this is not the case. This complicates the interpretation of the ancestral factor VIII gene haplotype(s) from which the haplotypes existing in present day populations must have evolved.

The differences in allele frequencies between populations have implications for linkage studies in haemophilia A: the loci of greatest diagnostic potential may differ according to the ethnicity of the family to be investigated and this has to be borne in mind in assessing the investigative strategy.

Linkage disequilibrium has not been investigated for all loci, although it has been shown for some. There is strong linkage disequilibrium between the intron 18 BclI, intron 7 G/A, intron 19 HindIII, and intron 25 BglI RFLPs.[59 63 64] The

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

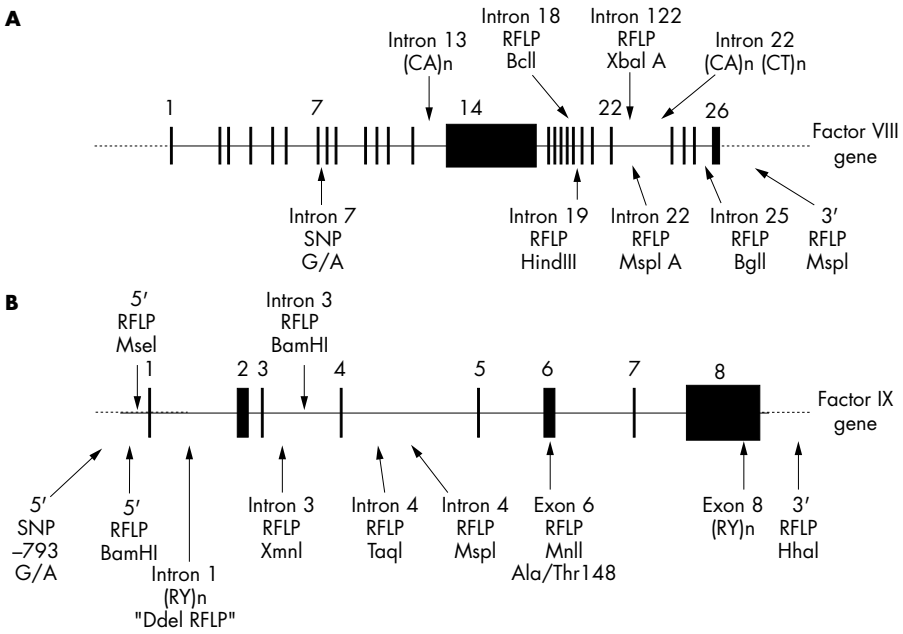


Figure 5 Some of the known polymorphisms in the human genes for (A) factor VIII and (B) factor IX. Source references: factor VIII gene intron 7 G/A,[37] intron 13 (CA)n,[38] intron 18 BclI,[39] intron 19 HindIII,[40] intron 22 XbaI A,[41 42] intron 22 MspI A,[43] intron 22 (CA)n,[44] intron 25 BglI,[45] 3′ MspI,[46] factor IX gene 5′ –793 G/A,[47] 5′ BamHI,[48] 5′ MseI,[49] intron 1 DdeI,[50] intron 3 XmnI,[50] intron 3 BamHI,[51] intron 4 TaqI,[52] intron 4 MspI,[53] exon 6 MnlI,[54] exon 8 (RY)n,[55] and 3′ HhaI.[56]

linkage disequilibrium in these cases is strong but not complete; therefore, additional informativity might be obtained from their inclusion in a linkage study. The intron 18 BclI and intron 22 XbaI A loci also show linkage disequilibrium[41]; however, this is less pronounced and their combined analysis provides much more informativity than either used on its own. For example, in whites, the informativity for combined use *v* individual use of XbaI A is approximately 65% *v* 49%,[41] in Chinese it is 52% *v* 49%,[61] and in Japanese it is 79% *v* 48%.[60]

XbaI A and MspI A are located within 737 bases of each other.[43] Despite their close proximity, the RFLPs are not in complete linkage disequilibrium. Haplotype analysis from a white population predicts an informativity of approximately 60% in linkage studies using both, compared with 47% using either on its own.[43]

Linkage disequilibrium, like allele frequencies, differs between ethnic groups and this should also be taken into account in assessing the linkage analysis strategy within a family.[34] Linkage disequilibrium and low rates of heterozgosity decrease the overall potential informativity of the polymorphisms in the factor VIII gene such that, in practice, about 10–15% of linkage studies are non-informative. The addition of further polymorphic loci to the small array already available would therefore be valuable.

FACTOR IX GENE POLYMORPHISMS

The human factor IX gene contains several SNPs, most of which fortuitously fall into the subcategory of RFLPs (fig 5B). The MnlI polymorphism[54] is located in exon 6 and brings about the amino acid substitution threonine/alanine at position 148 (known as the Malmö polymorphism[65]). Therefore, the factor IX protein itself is polymorphic and the Malmö polymorphism can be genotyped at the protein level using antibodies specific for each variant.[66 67] There are two microsatellite loci, one in intron 1[68] (the "DdeI" polymorphism[50]) and one in the 3′ untranslated region of exon 8[55] (fig 5B). These are of the type (RY)n (R, purine; Y, pyrimidine).

As seen for polymorphisms in the factor VIII gene, factor IX gene polymorphisms show ethnic variation, which in some

Table 3 Allele frequencies for four factor VIII gene polymorphisms frequently used in linkage studies in haemophilia A

			Frequency			
Location	Name	Alleles	Whites	Japanese	Chinese	American blacks
Intron 18	BclI	+	0.76	0.70	0.82	0.20
		–	0.24	0.30	0.18	0.80
Intron 19	HindIII	+	0.25		0.24	0.78
		–	0.75		0.76	0.22
Intron 22	XbaI A	+	0.58	0.59	0.57	
		–	0.42	0.41	0.43	
Intron 25	BglI	+	0.90	0.91	1	0.74
		–	0.10	0.09	0	0.26

Source references, white: BclI,[59] HindII,[59] XbaI A,[42] BglI,[45]; all Japanese data[60]; Chinese: BclI,[61] HindIII,[59] XbaI,[61]; American black: BclI,[59] HindII,[59] and BglI.[45]

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

Table 4 Allele frequencies for several factor IX gene polymorphisms in different ethnic groups

Location	Name	Alleles	Frequency			
			Whites	Japanese	Chinese	American blacks
5′	BamHI	+	0.02		0	0.36
		–	0.98		1	0.64
Intron 3	XmnI	+	0.29	0	0.01	0.12
		–	0.71	1	0.99	0.88
Intron 3	BamHI	+	0.06		0.02	0.13
		–	0.94		0.98	0.87
Intron 4	TaqI	+	0.35	0	0.01	0.14
		–	0.65	1	0.99	0.86
Intron 4	MspI	+	0.80	1	0	0.61
		–	0.20	0	1	0.39
Exon 6	MnlI	+	0.33		0.03	0.11
		–	0.67		0.97	0.89
3′	HhaI	+	0.61	0.24	0.17	0.57
		–	0.39	0.76	0.83	0.43

Source references: white data,[26] Japanese data,[69 70] Chinese data,[34 71 72] American black data.[34 72–74]

instances is quite pronounced (table 4). For example, the allele frequencies of the BamHI RFLPs in the 5′ region and in intron 3 are reasonable in the American black population, whereas these loci are essentially non-polymorphic in whites, Chinese, and Japanese (table 4). In contrast, the MnlI RFLP in exon 6 is common in whites and American blacks, but rare in Chinese (table 4). These differences in allele frequencies between racial groups impact upon linkage studies and influence which loci should be included for familial investigation in haemophilia B.[34]

The (RY)n microsatellites have been investigated in considerable depth.[55 68 73] The 3′ (exon 8) locus has the consensus $(GT)_a(ATGCGT)_4AG(AC)_bGCAT(AC)_3(AT)_2$ and four alleles have been demonstrated (I–IV) in which the values of "a" and "b" vary.[55] The intron 1 (RY)n motif comprises two units, A and B, with consensus sequence $GT(AC)_3(AT)_3(GT(AT)_1$ (I = 4 for the A unit and 5 for the B unit).[68] The polymorphism is caused by variability in the number of A and B units present. Both microsatellites show variation between racial groups.[55 73] Unlike the microsatellites within the factor VIII gene, the variability in the microsatellites of the factor IX gene does not appear to arise as a result of replication slippage. Instead, the patterns of variation observed are compatible with sister chromatid exchange (homologous recombination).[68]

There is strong circumstantial evidence that (RY)n sequences are involved in homologous recombination.[68] In this light, it is interesting to note that the BamHI RFLP which is 5′ to the intron 1 (RY)n is in linkage equilibrium with polymorphisms located between this (RY)n and the exon 8 (RY)n.[68] Similarly, the HhaI RFLP 3′ to the exon 8 (RY)n is in linkage equilibrium with polymorphisms between this (RY)n and the intron 1 (RY)n.[68] In contrast, the polymorphisms between the intron 1 and exon 8 (RY)n are in strong linkage disequilibrium.[26 50 72]

If the (RY)n sequences of the factor IX gene truly are involved in homologous recombination events, this has evident implications for linkage studies (the loss of linkage between polymorphic loci and the defect in the factor IX gene). However, this has not been seen, although the possibility of its occurrence argues for the inclusion of several loci across the factor IX gene in a linkage study.

Linkage disequilibrium has been investigated for some of the factor IX gene polymorphisms. In whites, the TaqI, XmnI, MspI, and MnlI RFLPs show pronounced linkage disequilibrium, such that the use of all four loci would only increase informativity to 55% compared with 45% when using the TaqI site alone.[26 50 72] The DdeI and HhaI polymorphisms, however, show much less disequilibrium,[72] possibly for the reasons discussed above.

In American blacks, the TaqI, MspI, and intron 3 BamHI loci show linkage disequilibrium[51]; however, the 5′ BamHI, intron 1 (RY)n, and XmnI polymorphisms appear to be in equilibrium.[51 72 73] Once again, these observations may be explained by the involvement of (RY)n sequences in homologous recombination.

> "As seen for polymorphisms in the factor VIII gene, factor IX gene polymorphisms show ethnic variation, which in some instances is quite pronounced"

The G/A SNP at –793 is only 104 bases from the MseI RFLP at –698.[49] Loci in such close proximity would be expected to show complete linkage disequilibrium; however, this is not the case. Their combined use provides an informativity of 60% compared with approximately 45% for either marker used alone.[49]

No linkage disequilibrium has been seen between the 3′ HhaI RFLP and other intragenic loci. Taken together with the possible loss of linkage between loci either side of the (RY)n sequences, a theoretical minimum linkage study with high diagnostic potential can be proposed. This would use HhaI 3′ to the exon 8 (RY)n, a second polymorphism 5′ to the intron 1 (RY)n, and a third polymorphism between the (RY)n loci. The second and third loci should be chosen according to racial background.

LINKAGE ANALYSIS IN HAEMOPHILIA A AND B

In essence, the practical approach involves the extraction of genomic DNA from anticoagulated whole blood obtained from relevant family members, amplification of DNA spanning the target polymorphic locus using the polymerase chain reaction (PCR), and then analysis of the PCR product using any of a variety of techniques to distinguish the alleles present[75] (for example, restriction digestion, single strand conformation polymorphism analysis, heteroduplex analysis,[64] PCR product length analysis).

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

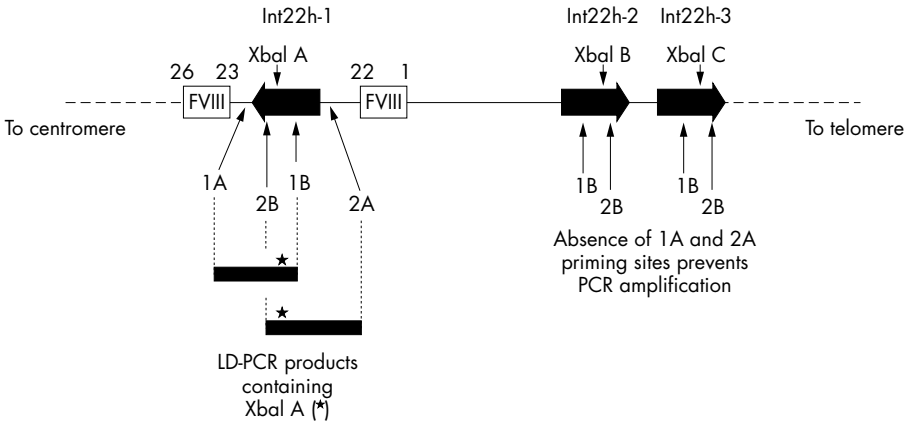


Figure 6 Genotyping the intron 22 XbaI A restriction fragment length polymorphism (RFLP) using long distance PCR (LD-PCR). A primer specific for factor VIII gene sequence (1A or 2A) is used in conjunction with a primer that flanks the XbaI A site and that hybridises within the int22h homologue (1B or 2B, respectively). The resulting LD-PCR product (~ 7 kb) contains XbaI A (asterisk) and can be genotyped by XbaI digestion. The analysis is specific for XbaI A because the sequences to which primers 1A or 2A hybridise do not occur at the extragenic int22h loci. Primers 1B or 2B can hybridise, but in the absence of primers 1A and 2A amplification cannot take place.

For most of the polymorphic loci in both genes this is a straightforward task and is the mainstream approach in linkage analyses.[34 75] However, for the XbaI A and MspI A RFLPs in intron 22 of the factor VIII gene, the task is complicated by the extragenic homologues (XbaI B and C, and MspI B and C). Until recently, specific amplification of these loci in int22h-1 was not possible. However, the introduction of long distance PCR (LD-PCR), in which stretches of genomic DNA as long as tens of kilobases can be amplified from small amounts of starting material, paved the way for the solution of specific amplification of the int22h-1 loci.[76 77]

The approach is outlined in fig 6. One LD-PCR primer hybridises to all three int22 homologues, whereas a second primer hybridises only to the factor VIII gene sequence outside the int22h-1 region. In the LD-PCR, exponential amplification is only possible for the int22h-1 target as a result of the hybridisation of both primers; int22h-2 and int22h-3 fail to amplify because only one primer can hybridise.

The resulting LD-PCR product can then be genotyped for XbaI A by digestion with this restriction endonuclease,[76 77] or it can be used as the substrate in a second round of PCR that contains primers for amplification of the MspI A site also contained within the LD-PCR product. MspI A can then be genotyped by digestion of the second round product with MspI.[77]

It is clear that PCR has provided a major step forward for the analysis of polymorphisms in the haemophilias. The traditional approach for RFLP analysis, using Southern blot, was labour intensive, time consuming, hazardous, and required relatively large quantities of high molecular weight DNA. In comparison, PCR is rapid, non-labourious, non-hazardous, and requires only small quantities of DNA, which can be degraded to a considerable degree.

Therefore, it is not surprising that the use of PCR is now common to most, if not all, laboratories engaged in genetic testing in the haemophilias. However, the methods used for post-PCR product analysis differ enormously. At one end of the spectrum there are low tech approaches, such as agarose gel electrophoresis followed by ethidium bromide staining, whereas at the other there are the high tech approaches such as the Light Cycler™ or fluorescence based capillary electrophoresis. The choice within laboratories reflects many parameters that need not concern us here; the important point is that post-PCR product analysis is flexible and amenable to the most up to date technologies in molecular biology. Whether "DNA chips"[78–80] (fig 7) may one day become useful for linkage studies in the haemophilias remains to be seen: the evolving movement towards direct mutation detection covered in the next section may ultimately result in linkage analysis being pushed into the sidelines.

THE MOLECULAR BASIS OF THE HAEMOPHILIAS

The mutations causing haemophilia A and B have been localised and characterised in several hundreds of patients. What is immediately evident from the enormous number of mutations that have been elucidated is that the molecular basis of the haemophilias is extremely diverse. This review will focus on some of the principal generalisations that emerge from an inspection of the vast databases of haemophilia mutations and it will consider some specific mutations that illustrate the differing molecular pathologies that exist.

The relation between cause and effect is not always clear, and only in a minority of mutations do we have a good understanding of how haemophilia ensues from the gene defect that is present. Unfortunately, this is mostly confined to those mutations whose effect is self evident in that the gene is disrupted, translation is prematurely terminated, activation of the protein is hindered, or mRNA splicing is corrupted, etc. For most mutations, there is no obvious explanation for effect, in which case one is left with candidate explanations such as protein destabilisation, incorrect folding, perturbation of structurally or functionally important regions, and so on. These are often difficult possibilities to prove. However, *x* ray crystallography and advances in software for molecular modelling and prediction of protein structure have permitted these possibilities to be explored for some mutations.[81–84]

OVERVIEW OF MUTATIONS

Mutations that have been characterised in the haemophilias are listed in the following databases on the World Wide Web (these databases are freely accessible).

- Haemophilia A: http://europium.csc.mrc.ac.uk/usr/WWW/WebPages/main.dir/main.htm (HAMSTeRS, The Haemophilia A Mutation, Structure, Test, Resource Site).
- Haemophilia B: http://www.umds.ac.uk/molgen/ (Haemophilia B Mutations Database).
- Both can be accessed via: http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html (Human Gene Mutation Database).

The databases are extensive listings that represent a vast coverage both of published and unpublished findings.[85] From these listings, several generalisations emerge that are common to both haemophilia A and B. The following is a synthesis of those generalisations.

TYPES OF MUTATION

Point mutations, deletions, insertions, and rearrangements/inversions are found in both the factor VIII and factor IX

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

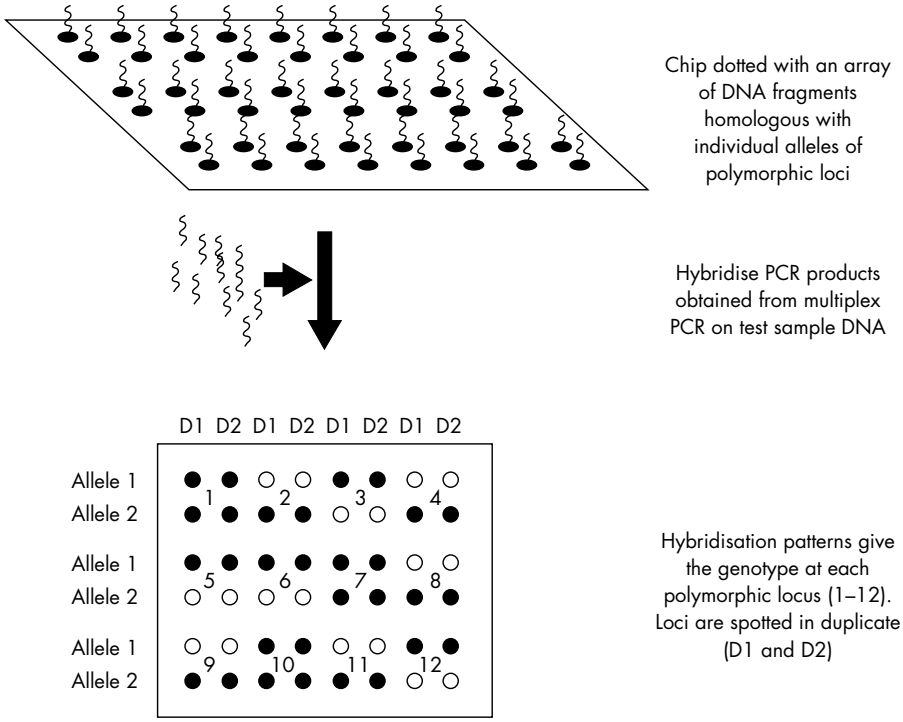


Figure 7 One approach to polymorphism screening using DNA chips. Multiplex PCR could be used to generate locus specific fragments for hybridisation to the chip or, conceivably, fragmented, labelled genomic DNA could be hybridised directly, without amplification.

genes. Point mutations (single nucleotide substitutions) are the most common gene defect and are present in approximately 90% of patients. Deletions are the second most common gene defect and are present in approximately 5–10% of patients. Insertions and rearrangements/inversions are quite rare within the haemophilia population, with the exception of an inversion that is prevalent among patients with severe haemophilia A (this inversion, the intron 22 inversion, will be discussed in more detail later).

Point mutations

The point mutations that occur in the haemophilias comprise missense mutations (these change a codon so that a different amino acid is encoded), nonsense point mutations (these change an amino acid codon into a translation stop codon), and mRNA splice site point mutations (these corrupt a true mRNA splice site, or create a novel one).

Missense mutations

The severity of haemophilia arising from missense point mutations depends very much on the nature of the amino acid substitution and its location. Semiconservative amino acid substitutions (the new amino acid is similar to the normal one either with respect to charge, hydrophobicity, polarity, or shape) tend to be associated with less severe disease except when they occur in regions of structural or functional importance. Non-conservative amino acid substitutions (the new amino acid is very dissimilar to the normal amino acid) tend to be associated with moderate to severe disease wherever they occur in the protein. The scope for structural perturbation/destabilisation is greater for non-conservative amino acid substitutions; therefore, there is a higher probability of their having a more detrimental overall effect.

Nonsense mutations

The outcome is expected to be a truncated protein molecule. This may, or may not, get as far as entering the circulation, but even if it does, truncation is highly likely to have made it defunct. It is therefore not surprising that nonsense mutations are associated with severe haemophilia. Exon skipping is a further possibility arising from a nonsense mutation[86 87] and this is also extremely detrimental: an in frame skip will result in a protein lacking the amino acids encoded by the skipped exon, an out of frame skip will result in a frame shift.

mRNA splice site mutations

Mutations that destroy or create mRNA splice sites are associated with variable severity of haemophilia; this very much depends on whether some correct transcripts can be processed (mild to moderate disease) or whether there is a complete loss of correct mRNA processing (severe disease). Exon skipping is a further possible consequence of splice site mutations, the outcome of which depends on whether the skip is in frame or results in a frame shift.[88]

Point mutations can be grouped according to whether or not they are classic CG site transitions. CG sites are hypermutable and give rise to C to T transitions or G to A transitions.[89] The mechanism underlying these transitions is understood[89] and is a major mechanism of human gene mutation. This is clearly exemplified in the haemophilias, in which approximately 30% of all distinct point mutations arise at CG sites. The remaining 70% of distinct point mutations do not occur at CG sites and may arise—for example, as a result of nucleotide misincorporation during DNA replication.

Point mutations may also be grouped according to whether or not they are recurrent. Recurrent mutations occur in both haemophilia A and B and are principally found at CG sites. The hypermutability of such sites underlies the independent recurrence of their transition events. There are rare examples of recurrent mutations that do not occur at CG sites. These are generally explained by a founder event that has multiplied in successive generations. Evidence for this may include confined geographical distribution and identical polymorphic haplotype in all affected genes, irrespective of geographical distribution.[90 91] Founder mutations typically cause mild/moderate disease; a mutation causing severe disease has a low probability of survival over many generations.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

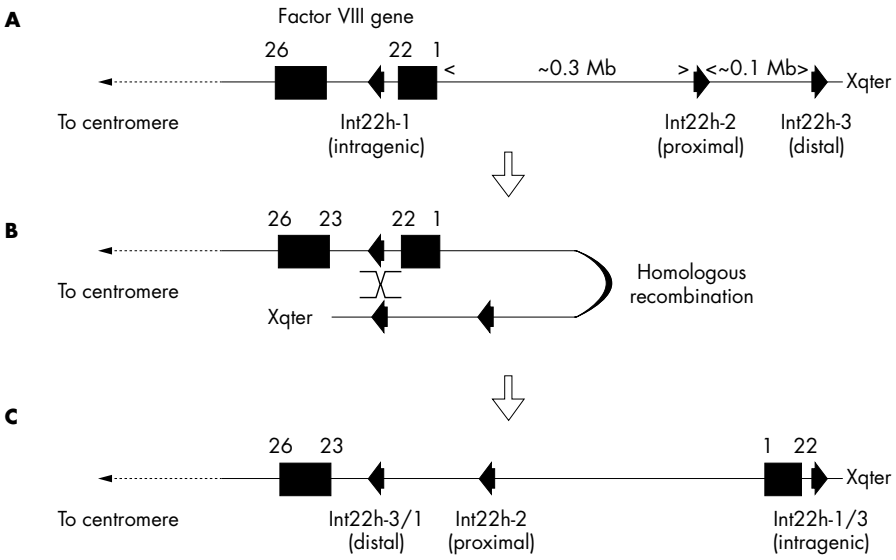


Figure 8 The intron 22 inversion of the factor VIII gene. (A) The normal configuration of the factor VIII gene and int22h homologues. (B) The distal or the proximal int22h homologue aligns with the intragenic int22h homologue (distal shown here). Homologous recombination takes place. (C) Linearisation of the X chromosome gives a completely disrupted factor VIII gene: exons 1 to 22 are displaced towards the telomere and are orientated in a direction opposite to their normal orientation. A contiguous factor VIII gene transcript running from exon 1 to 26 is no longer possible and severe haemophilia A results.

Deletions

Deletions of the factor VIII and factor IX gene include whole gene deletions, partial gene deletions at the 5′ end, 3′ end, or within the gene, and microdeletions of one to several base pairs (gross deletions are compiled in several reviews[14 28 92 93]). They do not appear to cluster to a specific region of either gene, but appear to be randomly distributed. A deletion has a high probability of destroying genetic function, removing a chunk of protein, or introducing a frame shift, all of which are extremely detrimental. Not surprisingly therefore, deletions are generally associated with severe disease.

Insertions

Insertions have been characterised in both the factor VIII and factor IX genes. Like deletions, these can be gross or as small as one or several nucleotides, and gene function or the gene product can be very adversely affected (gross insertions are compiled in several reviews[14 28 92 93]).

> "Deletions are generally associated with severe disease"

Rearrangements/inversions

These are rare gene defects in the haemophilias, except for an inversion involving intron 22 of the factor VIII gene, which causes severe haemophilia A, and which is found in 40–50% of patients with severe disease.[94 95]

The intron 22 inversion arises through homologous recombination between int22h-1 and int22h-2 (proximal) or int22h-3 (distal) during meiosis[94 95] (fig 8A–C). After homologous alignment, recombination results in the 5′ portion of the factor VIII gene being juxtaposed to the telomeric DNA, whereas the relevant extragenic homologue becomes juxtaposed to the 3′ portion of the factor VIII gene (fig 8B). In the re-linearised chromosome, the result is an inversion of the sequence between int22h-1 and int22h-2 or int22h-3, depending on which extragenic copy aligned (fig 8C). The factor VIII gene is completely disrupted: introns 1 to 22 are moved away from their normal context and their orientation is inverted. Severe haemophilia A results.

The intron 22 inversion is principally an error of DNA replication during spermatogenesis and not oogenesis[96]: in males, the absence of a second X chromosome for homologous alignment during meiosis would favour intrachromosomal alignments where they are possible. Thus, in haemophilia A families in which the intron 22 inversion is the causative gene defect, the defect can often be shown to have originated in an unaffected male relative, and in sporadic cases, this is often the patient's grandfather on his mother's side of the family.[96] The distal inversion is more common than the proximal inversion,[96] which may be explained by the greater genetic distance between the factor VIII gene and the distal int22h-3 homologue facilitating the formation of the loop required for alignment to take place.

The intron 22 inversion was, until recently, diagnosed using the labour intensive approach of Southern blot analysis (the DNA fragmentation pattern obtained for the normal gene differs from that obtained for an inverted gene, and the patterns for the distal and proximal inversions also differ).[94 95] This has now been superseded by an approach using LD-PCR (fig 9).[97] This impressive strategy detects normal and inverted genes; however, it does not distinguish proximal inversion from distal inversion.

Although most intron 22 inversions involve int22h-1 and 2/3, atypical inversions have been reported.[94 98 99] Detailed analysis of one of these revealed a third, truncated copy of int22h between the factor VIII gene and int22h-2, and this was involved in the inversion.[98] These atypical inversions give anomalous patterns using Southern blot. They may fail to be detected using LD-PCR and this has implications for diagnosis using this approach. Fortunately, they appear to be rare.

MOLECULAR PATHOLOGY: CAUSE–EFFECT RELATIONS

The following detailed examples offer a glimpse into the many and varied mechanisms through which mutations bring about their effects. Mutations can exert their effects at any stage from genetic function right through to protein function, and often they reveal important properties of these functions that might otherwise have escaped realisation. The investigation of cause–effect relations is therefore of considerable value with potentially high rewards.

Nonsense mutations at CGA (arginine) codons

CG transitions can give rise to missense or nonsense mutations, depending upon the codon in which the CG site

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

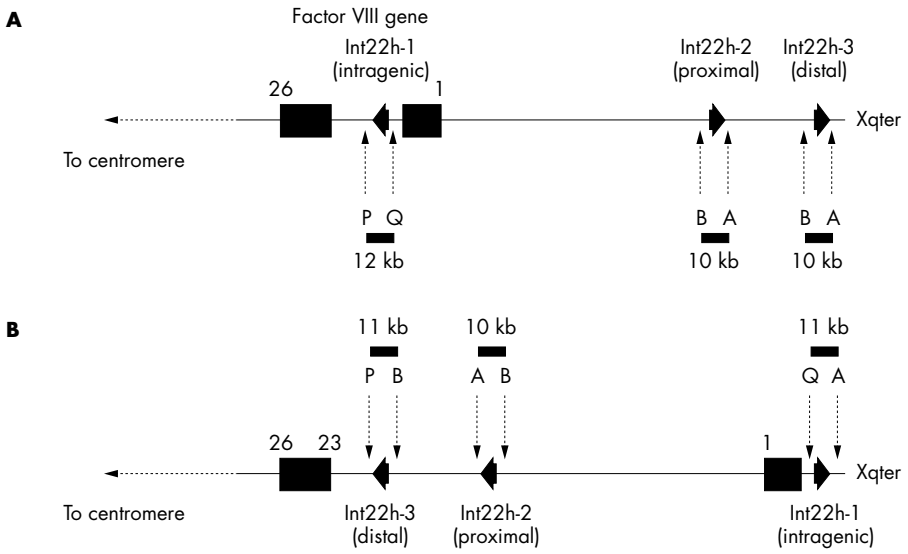


Figure 9 Detection of the intron 22 inversion of the factor VIII gene using long distance PCR (LD-PCR).[97] Primers P and Q hybridise specifically to sequences within the factor VIII gene and flank int22h-1. Primers A and B hybridise specifically to sequences flanking int22h-2 and int22h-3. The normal factor VIII gene therefore gives the LD-PCR products P + Q (12 kb) and A + B (10 kb). The intron 22 inversion results in the hybridisation site for primer Q being displaced towards the telomere, whereas the hybridisation site for primer B from the distal or proximal int22h homologue (depending on which was involved in the homologous recombination leading to the inversion, distal shown here) is brought into context with primer P. The inverted gene therefore gives the LD-PCR products P + B (11 kb), A + B (10 kb), and Q + A (11 kb). An A + B product is always obtained because one copy of the int22h homologues is unaffected by the inversion.

resides and whether C to T or G to A transition has occurred. C to T transition is particularly damaging at arginine codons of the type CGA: the result is the translation stop codon TGA (a nonsense mutation). As discussed above, nonsense mutations are particularly detrimental and are associated with severe disease. The factor VIII open reading frame has 70 CG sites, 12 of which reside in arginine codons of the type CGA (fig 10A). The factor IX open reading frame has 20 CG sites, six of which reside in CGA codons (fig 10B). C to T transition resulting in nonsense mutations has been reported at all of these codons and is associated with severe haemophilia, either haemophilia A[100–104] or B[105–107]. As would be expected from the hypermutability of CG sites, many of these nonsense mutations are recurrent among patients with haemophilia.

Missense mutations at activation cleavage sites

The proteolytic cleavages that activate factor VIII and factor IX occur after arginine residues located at positions 372 and 1689 in factor VIII (fig 3C), and 145 and 180 in factor IX (fig 4C). The codons for these arginine residues all contain the

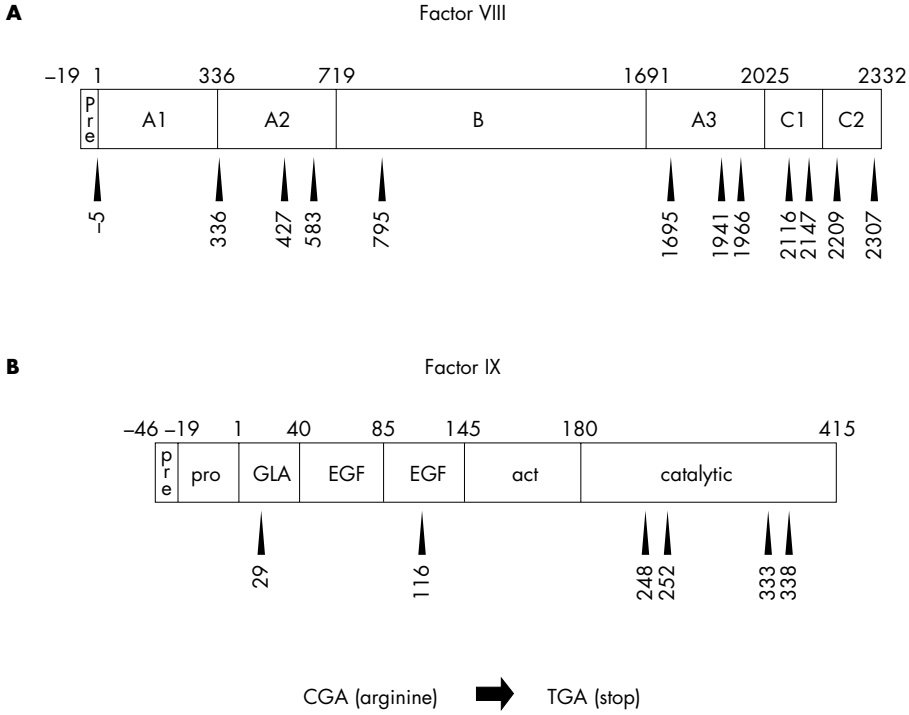


Figure 10 Arginine residues encoded by CGA within (A) the factor VIII and (B) the factor IX protein sequences. Nonsense mutations are brought about by C to T transition within the CG site of these codons. These are recurrent gene defects within haemophilia populations.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

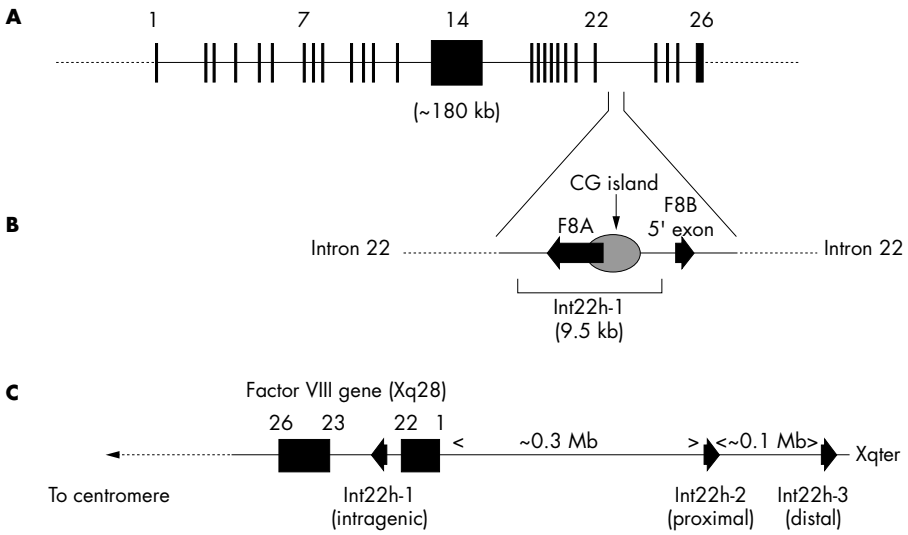


Figure 11 Mutations as a result of CG transitions at the codons encoding the activation cleavage site arginine residues of (A) factor VIII and (B) factor IX.

dinucleotide CG; transition of C to T or G to A within any of these codons causes haemophilia, either haemophilia A[103 104 108] or B.[107 109 110] Either transition results in a missense mutation that causes the relevant activation cleavage site arginine to be replaced by another amino acid (fig 11A, B). This compromises the activation cleavage at that site. Normal activation can therefore take place at only one of the two necessary sites and the result is a factor VIII or factor IX molecule with decreased function. Partial activity may remain, depending upon the activation cleavage site that is lost and the amino acid substitution. For example, factor IX cleaved at the β cleavage site alone retains partial activity, whereas factor IX cleaved at the α cleavage site alone is inactive; β cleavage is essential for unmasking the serine protease function of the catalytic domain. Thus, mutant proteins such as factor IX Chapel Hill[111] (histidine 145) retain some residual activity upon cleavage at the β bond, despite loss of cleavage at the α bond. The same is true for the cysteine 145 mutation (factor IX Albuquerque[112]). Mild haemophilia B results. In contrast, factor IX Hilo[113] (glutamine 180) cannot be activated at the β cleavage site, which is essential for the function of the enzyme, and severe haemophilia B results.

Mutations affecting factor VIII binding to von Willebrand factor

In plasma, factor VIII circulates bound to von Willebrand factor and this interaction stabilises factor VIII.[16 17] The regions of factor VIII that were found to be important for this interaction are the N-terminal part of the A3 domain, the C2 domain and, more recently, the C1 domain.[114–117] Mutations that interfere with factor VIII binding to von Willebrand factor illustrate how different mechanisms can affect the same property.

Within the A3 domain, sulphation of tyrosine 1680 is crucial for factor VIII interaction with von Willebrand factor.[118] Missense mutations resulting in the replacement of this tyrosine residue result in decreased factor VIII binding to von Willebrand factor[104 118 119] Similarly, retention of the tyrosine residue, but prevention of sulphation, also decreases factor VIII binding to von Willebrand factor.[120] Clearly, the loss of sulphation is crucial, whether or not this is accompanied by structural/folding changes brought about by a mutant amino acid substitution.

Several mutations have been identified within the C1 domain that decrease factor VIII binding to von Willebrand factor.[117 121] Some of these occur in residues that are conserved in the discoidin family of proteins; therefore, they are more likely to play a structural role in the C1 domain than to be involved directly in intermolecular interactions with von Willebrand factor. The structural perturbation that ensues compromises the interaction between the two proteins. Examples of such mutations are serine 2119 to tyrosine, asparagine 2129 to serine, arginine 2150 to histidine, and proline 2153 to glutamine.[117] However, none of these mutations sheds light on the mechanism by which C1 influences the affinity of factor VIII for von Willebrand factor: residues within the C1 domain may directly interact with von Willebrand factor residues; alternatively, mutations in the C1 domain may have an indirect effect through perturbation of folding or structure.

Besides mutations within the factor VIII gene that affect interaction with von Willebrand factor, there are also mutations of the von Willebrand factor gene that decrease the affinity of von Willebrand factor for factor VIII (von Willebrand disease type 2N).[122] Von Willebrand disease type 2N is typified by decreased factor VIII values, but shows an autosomal pattern of inheritance (the von Willebrand factor gene is located on chromosome 12 at 12p12–12pter). For obvious reasons, von Willebrand disease type 2N has also been referred to as pseudohaemophilia A.[123] There is an increasing list of mutations underlying von Willebrand disease type 2N. These are listed on the World Wide Web at URL http://mmg2.im.med.umich.edu/vWF/ (von Willebrand Factor (VWF) Database).

Mutations affecting factor VIII secretion

Factor VIII values may be decreased in the circulation by mutations that compromise its secretion. The haemophilia A missense mutations arginine 593 to cysteine and asparagine 618 to serine cause intracellular accumulation of factor VIII when expressed in cell culture.[124] This was concurrent with decreased factor VIII activity and antigen in the conditioned medium of transfected cells. Intracellular accumulation and decreased secretion of factor VIII possessing either of these mutations may explain the reduction of both factor VIII activity and antigen in plasma of patients with mild haemophilia A and these gene defects.

Mutations affecting factor IX binding to factor VIII

Initial evidence indicating that the protease domain of activated factor IX is involved in binding to activated factor VIII was provided by the inhibition of the interaction by a monoclonal antibody directed against the protease domain.[125] Subsequent antibody and mutation studies have revealed a complex story.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

The factor IX protease domain contains a Ca^{2+} binding site in which glutamic acid residue 245 participates in binding the metal ion.[126] Substitution of glutamic acid 245 by valine (a haemophilia B mutation) results in a factor IX molecule that is rapidly cleaved between arginine 318 and serine 319 by activated factor XI.[127] This variant factor IX has a 15 times lower affinity for factor VIII before cleavage at arginine 318 and a 120 times lower affinity after cleavage.[127] Currently, it is believed that binding of Ca^{2+} to the protease domain protects factor IX from proteolysis after arginine 318, and this stabilises a neighbouring surface exposed helix (residues 330–338), which binds factor VIII.[84 127]

Mutations of eight of the nine residues in this helix (leucine 330 to arginine 338) are reported in the haemophilia B population and they probably cause haemophilia, at least in part, by directly interfering with factor IX–factor VIII interaction. In contrast, the haemophilia B mutation glutamic acid 245 to valine abolishes Ca^{2+} binding and this has an indirect effect on factor VIII binding, through destabilisation and possible structural effects, as discussed above.[128]

Mutations causing factor IX deficiency during anticoagulant treatment

Bleeding complications arising from oral anticoagulant treatment are not uncommon. An unusually selective decrease in factor IX procoagulant activity (1–3% of normal) has been observed in some patients receiving coumarin.[129] Normal factor IX activity was recovered on withdrawal of the anticoagulant. Two different missense mutations at alanine −10 were found to be implicated: alanine to valine and alanine to threonine. Alanine −10 lies within the pro-peptide domain, at a position that is essential for binding of the carboxylase that modifies glutamic acid residues in the adjacent GLA domain (fig 4C).[130 131] This carboxylase requires vitamin K as cofactor.[132 133] Coumarin derivatives interrupt the vitamin K cycle and lead to rapid depletion of the cofactor form of the vitamin.[134] Carboxylase activity is consequently decreased and this, in conjunction with a decreased affinity of the carboxylase for the mutant pro-domain,[131 135] results in the loss of GLA carboxylation and hence of factor IX activity. Presumably, in the absence of coumarin, normal vitamin K concentrations supply the carboxylase adequately and this permits appropriate processing of the GLA domain, even though the affinity of the carboxylase for the pro-domain may be decreased by the alanine −10 mutations.

Mutations in the factor IX promoter and haemophilia B Leyden

The mutations discussed so far have affected either the factor VIII or factor IX coding sequence, with a consequent impact upon the translated protein. However, certain mutations within the promoter of the factor IX gene do not result in a qualitative change in the factor IX molecule. These mutations lie within sequences to which proteins involved in transcription bind. They disrupt or interfere with protein–DNA interactions in the promoter and so compromise the efficiency of transcription.

Some promoter mutations cause haemophilia B that is present throughout life.[136] However, others cause a variant of haemophilia B known as haemophilia B Leyden, which is characterised by severe factor IX deficiency at birth and through childhood, an increase in factor IX values from puberty through adolescence, and near normal factor IX values throughout adulthood.[137]

The molecular pathology of haemophilia B Leyden has been investigated intensively. The promoter of the factor IX gene contains a binding site for the CCAAT/enhancer binding protein[138] (C/EBP) and for the rat liver protein LF-A1/HNF4 (fig 12).[136] These proteins are involved in transcription; therefore, the binding sites may act as constitutive promoter

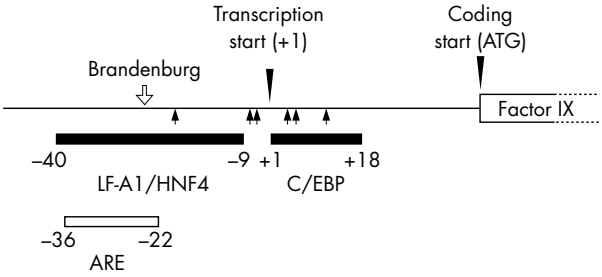


Figure 12 The promoter region of the human factor IX gene showing the androgen response element (ARE) and the regions implicated in binding LF-A1/HNF4, C/EBP. Coordinates are relative to the transcription start site (+1). Haemophilia B Leyden mutations (closed arrows) lie outside the ARE but within the LF-A1/HNF4 and C/EBP regions implicated in constitutive expression. In contrast, the promoter mutation G-26C (haemophilia B Brandenburg, open arrow) causes lifelong haemophilia B and lies within both the ARE and the −9 to −40 region associated with constitutive expression.

elements for gene expression. The promoter also contains a candidate androgen response element (ARE), which overlaps the LF-A1/HNF4 element (fig 12).[136 139] This may act as a hormonally regulated promoter element for gene expression.

> "Certain mutations within the promoter of the factor IX gene disrupt or interfere with protein–DNA interactions in the promoter and so compromise the efficiency of transcription"

Before puberty, factor IX gene expression is regulated by constitutive promoter elements (C/EBP, LF-A1/HNF4, and possibly others). Mutations within these elements disrupt binding of the relevant proteins for transcription and haemophilia B results. At puberty, and thereafter, it is assumed that testosterone dependent transcription, mediated by the ARE, comes into play. Thus, provided that a promoter mutation does not lie within the ARE, amelioration of the haemophilia B is possible after puberty.

In contrast, promoter mutations that cause life long haemophilia B, such as haemophilia B Brandenburg[136] (G-26C), have been found to lie within the region of overlap between the ARE and the LF-A1/HNF4 element and appear to interfere with both constitutive and hormonally regulated factor IX gene expression (fig 12).

Haemophilia A arising from mutations within the promoter of the factor VIII gene has not yet been demonstrated.

MUTATION DETECTION STRATEGIES

Studies aimed at defining the causative mutation within haemophilia cohorts have repeatedly shown that most patients possess a defect within the essential regions of the factor VIII or factor IX gene. These regions comprise the 5′ promoter sequence, the exons, the intron–exon boundaries, and the 3′ region around the polyadenylation signal sequence.

Therefore, the most direct strategy for mutation detection would be to amplify these regions from genomic DNA using PCR and screen the PCR products for a mutation. In the case of the factor IX gene, this is a straightforward task: the gene has only eight exons, the largest of which is less than 2 kb, so that fewer than 10 amplifications are required to cover the essential regions of the gene.[105]

In contrast, the large size and complexity of the factor VIII gene would necessitate almost 30 amplifications of genomic DNA to cover the essential regions. This is manageable and is the strategy used in some laboratories.[75 140–142] Alternatively, factor VIII mRNA can be used to screen exons 1 to 13 and 15 to 26.[143] The large exon 14 and the 5′ promoter and 3′ polyadenylation signal region can be screened from genomic DNA.[143]

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

Factor VIII mRNA is not obtained from the accepted primary site of transcription (the liver), but from circulating white blood cells in which the transcript is believed to be present as a result of ectopic (leaky) expression.[143]

Various mutation detection techniques can be used to screen the PCR products from the factor VIII or factor IX genes.[75] One of the most informative is chemical mismatch cleavage,[144] the approach used for the characterisation of haemophilia B mutations within the Swedish population[145] and within the UK population.[146] It has also been used extensively in the current effort to characterise the mutations within the UK haemophilia A population.[147] Solid phase, fluorescence based chemical mismatch cleavage has provided a rapid route to screening the essential regions of the complex factor VIII gene in large numbers of patients.[148]

Direct nucleotide sequence analysis using automatic DNA sequencers is becoming more mainstream[142] and confident results can be expected for male DNA (hemizygous) samples. Sequencing should be interpreted cautiously for female DNA because heterozygosity may fail to show in the sequencing data. Multiplex amplification of all of the essential regions of the factor IX gene in a single PCR, followed by sequencing,[149] represents a further step forward and one that could be applied to the factor VIII gene.

MUTATIONS AND THE RISK OF INHIBITORS

The production of neutralising antibodies in response to infused factor VIII or factor IX has always been of considerable interest, principally because it is a major complication of replacement treatment.[150 151] It is also potentially a major complication of gene therapy; therefore, insight into the factors that predispose towards inhibitor production would be of considerable value. Evidently, the mutation underlying the haemophilia is important[141 150–153]: mutations associated with the absence of a gene product, such as deletions or nonsense mutations, confer a high risk for inhibitor production; mutations associated with the presence of a gene product (even very low amounts of the protein) confer a low risk for inhibitor production. However, in reality, the situation is not so clear cut. Among patients with identical mutations, some may produce inhibitors and others may not.[141 154] Clearly other factors are implicated. The possibility that these may include the genotype at the (MHC) locus has been investigated in patients with severe haemophilia A, with and without the intron 22 inversion.[155 156] Only a weak association between human MHC (HLA) class II genotype and the development of inhibitor antibodies against factor VIII was obtained; this was slightly more pronounced in patients with the inversion than in those without.

Subtle effects may attenuate the relation between a haemophilia mutation and the risk of inhibitor production. For example, one would expect small insertions or deletions resulting in an alteration of the protein translation reading frame (frameshift mutations) to confer a high risk of inhibitor production. However, several groups have found the absence of inhibitors among patients with haemophilia A possessing a small deletion or insertion within a run of the same nucleotide.[141 157] A proposed explanation is that transcriptional or translational errors may lead, fortuitously, to the occasional correction of the reading frame and this may lead to sufficient endogenous factor VIII to explain immune tolerance.[141 157 158] Patients with a small deletion or insertion not within a run of the same nucleotide may not be so lucky.[141]

It therefore appears, at present, a difficult task, even with the knowledge of a patient's causative mutation, to predict an inhibitor response. Fortunately, however, inhibitors are a rare phenomenon among patients with haemophilia (apparently rarer in haemophilia B than haemophilia A). As one might expect, they are principally found within the severe group (high probability of the absence of a gene product), where they are found in 10–15% of individuals.[150 153]

FACTOR VIII AND FACTOR IX TURNOVER

The steady state concentrations of factor VIII and factor IX in the circulation reflect a balance of synthesis and removal. The clearance mechanism for both factors appears to be mediated, at least in part, by the low density lipoprotein receptor related protein (LRP), a liver multiligand endocytic receptor. Although the interaction of factor VIII with LRP is not disputed, there is some controversy concerning the region of factor VIII involved in binding and the effect of von Willebrand factor on the interaction. Factor VIII may bind LRP via a region within the light chain[159] and this may be inhibited by von Willebrand factor,[159] which also binds factor VIII via the light chain.[118] This could then explain the beneficial effect of von Willebrand factor on the in vivo survival of factor VIII.[16 17] Alternatively, factor VIII may bind LRP via the A2 domain in an interaction that is not influenced by von Willebrand factor.[160] In vivo studies in von Willebrand factor deficient mice suggest that the accelerated clearance of factor VIII seen in the absence of von Willebrand factor may be the result of LRP mediated factor VIII catabolism.[161] Clearly, LRP mediated removal of factor VIII appears to be real, and whether or not this is modulated by von Willebrand factor remains to be clarified.

LRP may also be involved in the catabolism of factor IX. Factor IX interaction with LRP appears to require activation of the coagulation factor by activated factor XI.[162] Delivery of activated factor IX to LRP appears to require proteoglycans on the cell surface.[162] In vitro, degradation of factor IX by cells deficient in either LRP or proteoglycans was significantly decreased compared with cells expressing these proteins. The relevance of these findings has yet to be established in vivo, although they suggest that one route to catabolism of activated factor IX may be via a proteoglycan on the cell surface, which delivers factor IX to LRP, which then targets the factor IX to the intracellular degradation pathway.

> "Insights gained from mutations and their effects—the molecular pathology of disease—have revealed both subtle and major complexities at the genetic and protein levels"

CONCLUDING COMMENTS

The exploration of the factor VIII and factor IX genes and their gene products over the past 15 years has provided enormous amounts of information and has expanded into many and varied territories. To review all of these comprehensively would require several volumes; therefore, this review has necessarily been restricted. The areas covered reveal the ingenuity of the research community (scientist and clinician alike) and humankind's incessant struggle for knowledge and understanding. As each question is answered, and new questions arise, the resolution of our understanding sharpens.

Insights gained from mutations and their effects—the molecular pathology of disease—have revealed both subtle and major complexities at the genetic and protein levels. Similarly, insights gained from polymorphisms and their ethnic variations hint at a complex evolution awaiting further exploration. Finally, insights gained from the investigation of patients' genetics and the aetiology of the inhibitor complication are paving the path for gene therapy in previously untreated patients. The molecular biology of the haemophilias has provided food for research and food for thought, and the feast seems certain to continue into the distant future.

APPENDIX A

DEFINITIONS OF POLYMORPHISMS FOUND IN THE FACTOR VIII AND FACTOR IX GENES, AND PROPERTIES RELEVANT TO LINKAGE STUDIES IN THE HAEMOPHILIAS

Single nucleotide polymorphisms

Single nucleotide polymorphisms (SNPs) are precisely what their name suggests: nucleotide positions at which different

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

nucleotides are present in the gene in the normal population. Typically, they are biallelic (one of two nucleotides is present within any given gene).

Many SNPs fortuitously reside within recognition sequences for restriction endonucleases and this provides a convenient means by which to genotype them. Such SNPs are better known as restriction fragment length polymorphisms (RFLPs).

SNPs can be intronic or exonic. Those within exons can either be silent in that they do not change the amino acid encoded by the codon in which they occur, or they can change a codon so that a different amino acid is specified. The latter give rise to amino acid polymorphisms within the protein sequence.

Microsatellites

Microsatellites are sequences built up from multiple repeats of a given sequence motif and are polymorphic in that the number of repeats present within the gene varies in the normal population. They are typically multiallelic. The human genome contains a diversity of microsatellites, just one example of which is the sequence (CA)n, where "n" represents a variable number of repeats among different genes in the normal population.

Frequency

The alleles of SNPs and microsatellites occur with specific frequencies in the general population. In the case of biallelic SNPs, the equal occurrence of both alleles would result in a relative frequency of 0.5 for each. However, in reality, such equivalence in frequency between both alleles is rarely seen and a preponderance of one allele over the other is more usual—for example, (0.1/0.9) or (0.4/0.6) (allele 1/allele 2).

In the case of microsatellites, several alleles are usually present within the general population and their relative frequencies are not equal. One, two, or three alleles may predominate and other alleles may be more rare. The existence of several alleles is an important property of microsatellites in the context of linkage analysis (see later).

The allele frequencies of any polymorphism may vary between people from different racial backgrounds and, indeed, some loci may be polymorphic within one racial group but non-polymorphic within another.

Heterozygosity, homozygosity, and hemizygosity

Females with a normal karyotype (46:XX) have two copies of each polymorphism located on the X chromosome and this, of course, includes those polymorphic loci found within the factor VIII and factor IX genes. For a given polymorphic locus within these genes, a female may have identical alleles (homozygous) or she may have one allelic variant in one gene and a different allelic variant in the other (heterozygous).

The probability of heterozygosity can be estimated from the allele frequencies in the general population. For example, a biallelic locus for which the alleles occur at frequencies of 0.1/0.9 can be expected to show the following distribution of genotypes:

Homozygous (allele 1/allele 1): $0.1 \times 0.1 = 0.01$.

Heterozygous (allele 1/allele 2): $0.1 \times 0.9 = 0.09$.

Heterozygous (allele 2/allele 1): $0.1 \times 0.9 = 0.09$.

Homozygous (allele 2/allele 2): $0.9 \times 0.9 = 0.81$.

Thus, the expected frequency of heterozygosity (the rate of heterozygosity) is $0.09 + 0.09 = 0.18$. Another way of interpreting this is that 18% of females may be expected to be heterozygous at that locus. An identical calculation done using a biallelic locus with allele frequencies of 0.5/0.5 yields a rate of heterozygosity of 0.5. This represents the maximum possible rate of heterozygosity for a biallelic locus and it can be interpreted to mean that 50% of females can be expected to be heterozygous at that locus.

Microsatellites yield much higher rates of heterozygosity because there are greater numbers of allelic variants and co-inheritance of identical alleles is intrinsically less probable than for a biallelic locus. For example, in a microsatellite with three allelic variants showing frequencies of 0.1 (allele 1), 0.6 (allele 2), and 0.3 (allele 3), the expected distribution of genotypes is as follows:

Homozygous (allele 1/allele 1): $0.1 \times 0.1 = 0.01$.

Homozygous (allele 2/allele 2): $0.6 \times 0.6 = 0.36$.

Homozygous (allele 3/allele 3): $0.3 \times 0.3 = 0.09$.

Heterozygous (allele 1/allele 2): $0.1 \times 0.6 = 0.06$.

Heterozygous (allele 2/allele 1): $0.6 \times 0.1 = 0.06$.

Heterozygous (allele 1/allele 3): $0.1 \times 0.3 = 0.03$.

Heterozygous (allele 3/allele 1): $0.3 \times 0.1 = 0.03$.

Heterozygous (allele 2/allele 3): $0.6 \times 0.3 = 0.18$.

Heterozygous (allele 3/allele 2): $0.3 \times 0.6 = 0.18$.

From these calculations, the expected rate of heterozygosity is $0.06 + 0.06 + 0.03 + 0.03 + 0.18 + 0.18 = 0.54$. This exceeds the 0.5 maximum for biallelic loci and yet it has only involved a locus with three alleles, one of which is of very low frequency (allele 1, 0.1). In practice, microsatellites have many more than three alleles and it is not unusual to observe rates of heterozygosity approaching 0.8.

A male with a normal karyotype (46:XY) only has one factor VIII gene and factor IX gene, therefore he will have only one copy of a given polymorphic locus and only one allele will be present (hemizygous).

Informativity and non-informativity

The outcome of a linkage study cannot be predicted at the outset: the data must be generated and then interpreted to reveal whether the polymorphic loci investigated have permitted a defective (or normal) gene to be tracked through the family. Any polymorphic loci that permit gene tracking from one individual to another within a family are said to be informative for that individual. Conversely, those loci that do not permit the gene to be tracked from one individual to another are non-informative for that individual.

Evidently, heterozygosity is associated with informativity: the two genes can be discerned by virtue of their different polymorphic alleles and this permits them to be tracked separately. Homozygosity is associated with non-informativity because the two genes cannot be distinguished. However, homozygosity can be helpful in establishing haplotypes from genotypes (see below).

Genotypes and haplotypes

Analysis of polymorphic loci to determine the alleles present generates a genotype for each individual in a family. This must then be resolved into haplotypes (the alleles in one gene separated from those in the other). In the case of the haemophilias, this is made easier by virtue of hemizygosity in males: the genotype is the haplotype in the case of hemizygosity. This may then permit haplotypes in females to be established by subtraction of a known haplotype from the genotype to give the second haplotype.

Homozygosity in females also permits the determination of haplotypes: if a female is homozygous, both genes must have identical haplotypes for the polymorphic loci examined, and this may permit further haplotypes to be resolved within the family.

In phase

The haplotype found in the factor VIII or factor IX gene of the male haemophiliac within a family is said to be in phase with haemophilia: it identifies the defective factor VIII or factor IX gene. It is worth stressing that this haplotype undoubtedly exists in the normal population where it is not associated with a defective gene, it just so happens that, in the family being

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

investigated, the defective gene has this haplotype. Thus, discovery of the same haplotype in other members of the family does not necessarily define the presence of the defective gene. Inheritance of the defective gene has to be shown directly by analysis of relevant family members.

Linkage disequilibrium

Polymorphisms within a given gene are necessarily linked. When a certain allele at one locus always occurs in conjunction with a certain allele at another locus, the loci are said to show complete linkage disequilibrium. When the alleles at one locus occur entirely independently of the alleles at a second locus, the loci are said to show complete linkage equilibrium. Between these two extremes there is a continuous scale of linkage disequilibrium.

If two loci show complete linkage disequilibrium, the information obtained at one will be identical to that obtained at the other and no additional information will be gained by examination of both. If one locus is homozygous and non-informative, the other one will be also, so that informativity is not increased by investigation of both loci. For this reason it may seem redundant to investigate loci showing linkage disequilibrium. However, examination of both does permit corroboration of the information obtained from either. This can be valuable, lending confidence to the data generated in the study. Loci that are in linkage disequilibrium may therefore not increase the informativity of a study very much, but they do provide a double check of the data, and this may be invaluable.

Implications of these properties of polymorphisms

The goal of a linkage study in the haemophilias is to establish carrier status and/or offer prenatal diagnosis with confidence. This requires an informative genetic result, and this should be derived preferably from more than one informative polymorphic locus. For this to be achieved, the best loci to examine are those with the highest rates of heterozygosity and the lowest degree of linkage disequilibrium.

APPENDIX B

DATABASE ENTRIES AT THE NATIONAL CENTRE FOR BIOTECHNOLOGY INFORMATION (URL HTTP://WWW.NCBI.NLM.NIH.GOV/)

Factor VIII (F8C)

- Gene: NT_025284
- mRNA: NM_000132
- Protein: NP_000123
- Locus Link: LocusID 2157

F8A

- Gene: NT_025307
- mRNA: NM_012151
- Protein: NP_036283
- Locus Link: LocusID 8263

F8B

- Gene: NT_025284
- mRNA: M90707
- Protein: AAA58466
- Locus Link: LocusID 2157

F8 Int22h-1

- DNA: X86011 and X86012

Factor IX (F9)

- Gene: K02402
- mRNA: NM_000133
- Protein: NP_000124
- Locus Link: LocusID 2158

REFERENCES

1 **Gitschier J**, Wood WI, Goralka TM, *et al*. Characterization of the human factor VIII gene. *Nature* 1984;**312**:326–30.
2 **Yoshitake S**, Schach BG, Foster DC, *et al*. Nucleotide sequence of the gene for human factor IX (antihemophilic factor B). *Biochemistry* 1985;**24**:3736–50.
3 **Drayna D**, White R. The genetic linkage map of the human X chromosome. *Science* 1985;**230**:753–8.
4 **Patterson M**, Schwartz C, Bell M, *et al*. Physical mapping studies on the human X chromosome in the region Xq27–Xqter. *Genomics* 1987;**1**:297–306.
5 **Wollina K**, Bowen DJ, Syrbe G, *et al*. Female twins with severe Christmas disease (hemophilia B). *Thromb Haemost* 1993;**70**:774–6.
6 **Ingerslev J**, Schwartz M, Lamm LU, *et al*. Female haemophilia A in a family with seeming extreme bidirectional lyonization tendency: abnormal premature X-chromosome inactivation? *Clin Genet* 1989;**35**:41–8.
7 **Orstavik KH**, Scheibel E, Ingerslev J, *et al*. Absence of correlation between X chromosome inactivation pattern and plasma concentration of factor VIII and factor IX in carriers of haemophilia A and B. *Thromb Haemost* 2000;**83**:433–7.
8 **Levinson B**, Kenwrick S, Lakich D, *et al*. A transcribed gene in an intron of the human factor VIII gene. *Genomics* 1990;**7**:1–11.
9 **Levinson B**, Kenwrick S, Gamel P *et al*. Evidence for a third transcript from the human factor VIII gene. *Genomics* 1992;**14**:585–9.
10 **Naylor JA**, Buck D, Green P, *et al*. Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Hum Mol Genet* 1995;**4**:1217–24.
11 **Levinson B**, Bermingham JR, Jr, Metzenberg A, *et al*. Sequence of the human factor VIII-associated gene is conserved in mouse. *Genomics* 1992;**13**:862–5.
12 **Wood WI**, Capon DJ, Simonsen CC, *et al*. Expression of active human factor VIII from recombinant DNA clones. *Nature* 1984;**312**:330–7.
13 **Vehar GA**, Keyt B, Eaton D, *et al*. Structure of human factor VIII. *Nature* 1984;**312**:337–42.
14 **Tuddenham EGD**, Cooper DN. Factor VIII and haemophilia A. *Oxford Monographs on Medical Genetics* 1994;**25**:19–76.
15 **Eaton D**, Rodriguez H, Vehar GA. Proteolytic processing of human factor VIII. Correlation of specific cleavages by thrombin, factor Xa, and activated protein C with activation and inactivation of factor VIII coagulant activity. *Biochemistry* 1986;**25**:505–12.
16 **Weiss HJ**, Sussman II, Hoyer LW. Stabilization of factor VIII in plasma by the von Willebrand factor. Studies on posttransfusion and dissociated factor VIII and in patients with von Willebrand's disease. *J Clin Invest* 1977;**60**:390–404.
17 **Tuddenham EG**, Lane RS, Rotblat F, *et al*. Response to infusions of polyelectrolyte fractionated human factor VIII concentrate in human haemophilia A and von Willebrand's disease. *Br J Haematol* 1982;**52**:259–67.
18 **Hamer RJ**, Kodam JA, Beeser-Viser NH, *et al*. The effect of thrombin on the complex between factor VIII and von Willebrand factor. *Eur J Biochem* 1987;**167**:253–9.
19 **Saenko EL**, Shima M, Sarafanov AG. Role of activation of the coagulation factor VIII in interaction with vWf, phospholipid, and functioning within the factor Xase complex. *Trends Cardiovasc Med* 1999;**9**:185–92.
20 **Anson DS**, Choo KH, Rees DJ, *et al*. The gene structure of human anti-haemophilic factor IX. *EMBO J* 1984;**3**:1053–60.
21 **Di Scipio RG**, Kurachi K, Davie EW. Activation of human factor IX (Christmas factor). *J Clin Invest* 1978;**61**:1528–38.
22 **Walsh PN**, Bradford H, Sinha D, *et al*. Kinetics of the factor XIa catalyzed activation of human blood coagulation factor IX. *J Clin Invest* 1984;**73**:1392–9.
23 **Jesty J**, Morrison SA. The activation of factor IX by tissue factor–factor VII in a bovine plasma system lacking factor X. *Thromb Res* 1983;**32**:171–81.
24 **Osterud B**, Rapaport SI. Activation of factor IX by the reaction product of tissue factor and factor VII: additional pathway for initiating blood coagulation. *Proc Natl Acad Sci U S A* 1977;**74**:5260–4.
25 **Foster PA**, Zimmerman TS. Factor VIII structure and function. *Blood Rev* 1989;**3**:180–91.
26 **Giannelli F**. Factor IX. *Baillieres Clin Haematol* 1989;**2**:821–48.
27 **Lenting PJ**, van Mourik JA, Mertens K. The life cycle of coagulation factor VIII in view of its structure and function. *Blood* 1998;**92**:3983–96.
28 **Giannelli F**, Green PM. The molecular basis of haemophilia A and B. *Baillieres Clin Haematol* 1996; **9**:211–28.
29 **Jeffreys AJ**, Wilson V, Thein SL. Hypervariable "minisatellite" regions in human DNA. *Nature* 1985;**314**:67–73.
30 **Gill P**, Jeffreys AJ, Werrett DJ. Forensic application of DNA "fingerprints". *Nature* 1985;**318**:577–9.
31 **Paabo S**. Human evolution. *Trends Cell Biol* 1999;**9**:M13–16.
32 **Foley R**. The context of human genetic evolution. *Genome Res* 1998;**8**:339–47.
33 **Mibashan RS**, Peake IR, Nicolaides KH. Prenatal diagnosis of hemostatic disorders. *Methods Haematol* 1989;**21**:64–107.
34 **Peake IR**, Lillicrap DP, Boulyjenkov V, *et al*. Report of a joint WHO/WFH meeting on the control of haemophilia: carrier detection and prenatal diagnosis. *Blood Coagul Fibrinolysis* 1993;**4**:313–44.
35 **Venter JC**, Adams MD, Myers EW, *et al*. The sequence of the human genome. *Science* 2001;**291**:1304–51.
36 **Lander ES**, Linton LM, Birren B, *et al*. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–21.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

37 **Kogan S**, Gitschier J. Mutations and a polymorphism in the factor VIII gene discovered by denaturing gradient gel electrophoresis. *Proc Natl Acad Sci U S A* 1990;**87**:2092–6.
38 **Lalloz MR**, McVey JH, Pattinson JK, *et al*. Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet* 1991;**338**:207–11.
39 **Gitschier J**, Drayna D, Tuddenham EG, *et al*. Genetic mapping and diagnosis of haemophilia A achieved through a BclI polymorphism in the factor VIII gene. *Nature* 1985;**314**:738–40.
40 **Ahrens P**, Kruse TA, Schwartz M, *et al*. A new HindIII restriction fragment length polymorphism in the hemophilia A locus. *Hum Genet* 1987;**76**:127–8.
41 **Wion KL**, Tuddenham EG, Lawn RM. A new polymorphism in the factor VIII gene for prenatal diagnosis of hemophilia A. *Nucleic Acids Res* 1986;**14**:4535–42.
42 **Chan V**, Tong TM, Chan TP, *et al*. Multiple XbaI polymorphisms for carrier detection and prenatal diagnosis of haemophilia A. *Br J Haematol* 1989;**73**:497–500.
43 **Bowen DJ**, De Brasi CD, Larripa IB, *et al*. A new polymorphism in the human factor VIII gene: implications for linkage analysis in haemophilia A and for the evolution of int22h sequences. *Br J Haematol* 2000;**111**:544–8.
44 **Lalloz MRA**, McVey JH, Michaelides K, *et al*. Haemophilia A diagnosis by analysis of a novel dinucleotide tandem repeat sequence within the factor VIII gene. *Br J Haematol* 1992;**80**:3a.
45 **Antonarakis SE**, Waber PG, Kittur SD, *et al*. Hemophilia A. Detection of molecular defects and of carriers by DNA analysis. *N Engl J Med* 1985;**313**:842–8.
46 **Youssoufian H**, Phillips DG, Kazazian HH, Jr, *et al*. MspI polymorphism in the 3′ flanking region of the human factor VIII gene. *Nucleic Acids Res* 1987;**15**:6312.
47 **Rai HK**, Winship PR. A –793 G to A transition in the factor IX gene promoter is polymorphic in the Caucasian population. *Br J Haematol* 1996;**92**:501–3.
48 **Hay CW**, Robertson KA, Yong SL, *et al*. Use of a BamHI polymorphism in the factor IX gene for the determination of hemophilia B carrier status. *Blood* 1986;**67**:1508–11.
49 **Winship PR**, Nichols CE, Chuansumrit A, *et al*. An MseI RFLP in the 5′ flanking region of the factor IX gene: its use for haemophilia B carrier detection in Caucasian and Thai populations. *Br J Haematol* 1993;**84**:101–5.
50 **Winship PR**, Anson DS, Rizza CR, *et al*. Carrier detection in haemophilia B using two further intragenic restriction fragment length polymorphisms. *Nucleic Acids Res* 1984;**12**:8861–72.
51 **Driscoll MC**, Dispenzieri A, Tobias E, *et al*. A second BamHI DNA polymorphism and haplotype association in the factor IX gene. *Blood* 1988;**72**:61–5.
52 **Giannelli F**, Anson DS, Choo KH, *et al*. Characterisation and use of an intragenic polymorphic marker for detection of carriers of haemophilia B (factor IX deficiency). *Lancet* 1984;**1**:239–41.
53 **Camerino G**, Oberle I, Drayna D, *et al*. A new MspI restriction fragment length polymorphism in the hemophilia B locus. *Hum Genet* 1985;**71**:79–81.
54 **Graham JB**, Kunkel GR, Tennyson GS, *et al*. The Malmo polymorphism of factor IX: establishing the genotypes by rapid analysis of DNA. *Blood* 1989;**73**:2104–7.
55 **Sommer SS**, Tillotson VL, Vielhaber EL, *et al*. "Cryptic" dinucleotide polymorphism in the 3′ region of the factor IX gene shows substantial variation among different populations. *Hum Genet* 1994;**93**:357–8.
56 **Winship PR**, Rees DJ, Alkan M. Detection of polymorphisms at cytosine phosphoguanadine dinucleotides and diagnosis of haemophilia B carriers. *Lancet* 1989;**1**:631–4.
57 **Rubinsztein DC**, Amos W, Leggo J, *et al*. Microsatellite evolution—evidence for directionality and variation in rate between species. *Nat Genet* 1995;**10**:337–43.
58 **Jeffreys AJ**, Royle NJ, Wilson V, *et al*. Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* 1988;**332**:278–81.
59 **Graham JB**, Kunkel GR, Fowlkes DM, *et al*. The utility of a HindIII polymorphism of factor VIII examined by rapid DNA analysis. *Br J Haematol* 1990;**76**:75–9.
60 **Suehiro K**, Tanimoto M, Hamaguchi M, *et al*. Carrier detection in Japanese hemophilia A by use of three intragenic and two extragenic factor VIII DNA probes: a study of 24 kindreds. *J Lab Clin Med* 1988;**112**:314–18.
61 **Chan V**, Chan TK, Liu VW, *et al*. Restriction fragment length polymorphisms associated with factor VIII:C gene in Chinese. *Hum Genet* 1988;**79**:128–31.
62 **Bowen DJ**. An effective route for the isolation of human genetic loci exemplified using the BglI restriction fragment length polymorphism of the factor VIII gene. *Thromb Haemost* 2000;**83**:512–13.
63 **Janco RL**, Phillips JA, III, Orlando P, *et al*. Carrier testing strategy in haemophilia A. *Lancet* 1986;**1**:148–9.
64 **Howarth A**, Bowen DJ. Linkage analysis in haemophilia A: simultaneous genotyping of two polymorphisms of the human factor VIII gene using induced heteroduplex formation. *Haemophilia* 1998;**4**:812–19.
65 **McGraw RA**, Davis LM, Noyes CM, *et al*. Evidence for a prevalent dimorphism in the activation peptide of human coagulation factor IX. *Proc Natl Acad Sci U S A* 1985;**82**:2847–51.
66 **Smith KJ**, Thompson AR, McMullen BA, *et al*. Carrier testing in hemophilia B with an immunoassay that distinguishes a prevalent factor IX dimorphism. *Blood* 1987;**70**:1006–13.
67 **Wallmark A**, Ljung R, Nilsson IM. Determination of factor IX allotypes for carrier identification in haemophilia B. *Br J Haematol* 1987;**67**:427–32.
68 **Sarkar G**, Paynton C, Sommer SS. Segments containing alternating purine and pyrimidine dinucleotides: patterns of polymorphism in humans and prevalence throughout phylogeny. *Nucleic Acids Res* 1991;**19**:631–6.
69 **Kojima T**, Tanimoto M, Kamiya T, *et al*. Possible absence of common polymorphisms in coagulation factor IX gene in Japanese subjects. *Blood* 1987;**69**:349–52.
70 **Toyozumi H**, Kojima T, Matsushita T, *et al*. Diagnosis of hemophilia B carriers using two novel dinucleotide polymorphisms and Hha I RFLP of the factor IX gene in Japanese subjects. *Thromb Haemost* 1995;**74**:1009–14.
71 **de la SC**, Wu Q, Baas MJ, *et al*. Common intragenic and extragenic polymorphisms of blood coagulation factors VIII and IX are different in Chinese and Caucasian populations. *Clin Genet* 1990;**38**:434–40.
72 **Graham JB**, Kunkel GR, Egilmez NK, *et al*. The varying frequencies of five DNA polymorphisms of X-linked coagulant factor IX in eight ethnic groups. *Am J Hum Genet* 1991;**49**:537–44.
73 **Figueiredo MS**, Bowen DJ, Silva Junior WA, *et al*. Factor IX gene haplotypes in Brazilian blacks and characterization of unusual DdeI alleles. *Br J Haematol* 1994;**87**:789–96.
74 **Zhang M**, Chen SH, Scott CR, *et al*. The factor IX BamHI polymorphism: T-to-G transversion at the nucleotide sequence –561. The BamHI/MSPI haplotypes in blacks and Caucasians. *Hum Genet* 1989;**82**:283–4.
75 **Goodeve AC**. Laboratory methods for the genetic diagnosis of bleeding disorders. *Clin Lab Haematol* 1998;**20**:3–19.
76 **El Maarri O**, Oldenburg J, Caglayan SH. Intron 22-specific long PCR for the Xba I polymorphism in the factor VIII gene. *Br J Haematol* 1999;**105**:1120–2.
77 **De Brasi CD**, Bowen DJ, Collins PW *et al*. Specific analysis of the intron 22 XbaI polymorphism of the human factor VIII gene using long-distance PCR. *Br J Haematol* 1999;**107**:566–8.
78 **Hacia JG**, Fan JB, Ryder O, *et al*. Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat Genet* 1999;**22**:164–7.
79 **Wilgenbus KK**, Lichter P. DNA chip technology ante portas. *J Mol Med* 1999;**77**:761–8.
80 **Diamandis EP**. Sequencing with microarray technology—a powerful new tool for molecular diagnostics. *Clin Chem* 2000;**46**:1523–5.
81 **Pemberton S**, Lindley P, Zaitsev V, *et al*. A molecular model for the triplicated A domains of human factor VIII based on the crystal structure of human ceruloplasmin. *Blood* 1997;**89**:2413–21.
82 **Brandstetter H**, Bauer M, Huber R, *et al*. X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and hemophilia B. *Proc Natl Acad Sci U S A* 1995;**92**:9796–800.
83 **Mertens K**, Celie PH, Kolkman JA, *et al*. Factor VIII–factor IX interactions: molecular sites involved in enzyme–cofactor complex assembly. *Thromb Haemost* 1999;**82**:209–17.
84 **Bajaj SP**. Region of factor IXa protease domain that interacts with factor VIIIa: analysis of select hemophilia B mutants. *Thromb Haemost* 1999;**82**:218–25.
85 **Wacey AI**, Tuddenham EG. Mutation databases on the Web. *J Med Genet* 1998;**35**:529–33.
86 **Dietz HC**, Valle D, Francomano CA, *et al*. The skipping of constitutive exons in vivo induced by nonsense mutations. *Science* 1993;**259**:680–3.
87 **Liu HX**, Cartegni L, Zhang MQ, *et al*. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat Genet* 2001;**27**:55–8.
88 **Tavassoli K**, Eigel A, Wilke K, *et al*. Molecular diagnostics of 15 hemophilia A patients: characterization of eight novel mutations in the factor VIII gene, two of which result in exon skipping. *Hum Mutat* 1998;**12**:301–3.
89 **Duncan BK**, Miller JH. Mutagenic deamination of cytosine residues in DNA. *Nature* 1980;**287**:560–1.
90 **Green PM**, Montandon AJ, Ljung R, *et al*. Haplotype analysis of identical factor IX mutants using PCR. *Thromb Haemost* 1992;**67**:66–9.
91 **Thompson AR**, Bajaj SP, Chen SH, *et al*. "Founder" effect in different families with haemophilia B mutation. *Lancet* 1990;**335**:418.
92 **Tuddenham EGD**, Cooper DN. Factor IX and haemophilia B. *Oxford Monographs on Medical Genetics* 1994;**25**:78–111.
93 **Antonarakis SE**. Molecular genetics of coagulation factor VIII gene and hemophilia A. *Thromb Haemost* 1995;**74**:322–8.
94 **Lakich D**, Kazazian HH, Jr, Antonarakis SE, *et al*. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nat Genet* 1993;**5**:236–41.
95 **Naylor J**, Brinke A, Hassock S, *et al*. Characteristic mRNA abnormality found in half the patients with severe haemophilia A is due to large DNA inversions. *Hum Mol Genet* 1993;**2**:1773–8.
96 **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**:2206–12.
97 **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**:1458–9.
98 **Naylor JA**, Nicholson P, Goodeve A, *et al*. A novel DNA inversion causing severe hemophilia A. *Blood* 1996;**87**:3255–61.
99 **Yamazaki E**, Mohri H, Inaba H, *et al*. Variant of intron 22 inversions in the factor VIII gene in severe hemophilia A. *Blood Coagul Fibrinolysis* 1997;**8**:445–9.
100 **Youssoufian H**, Kazazian HH, Jr, Phillips DG, *et al*. Recurrent mutations in haemophilia A give evidence for CpG mutation hotspots. *Nature* 1986;**324**:380–2.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

101 **Youssoufian H**, Antonarakis SE, Bell W, *et al.* Nonsense and missense mutations in hemophilia A: estimate of the relative mutation rate at CG dinucleotides. *Am J Hum Genet* 1988;**42**:718–25.
102 **Gitschier J**, Kogan S, Levinson B, *et al.* Mutations of factor VIII cleavage sites in hemophilia A. *Blood* 1988;**72**:1022–8.
103 **Pattinson JK**, Millar DS, McVey JH, *et al.* The molecular genetic analysis of hemophilia A: a directed search strategy for the detection of point mutations in the human factor VIII gene. *Blood* 1990;**76**:2242–8.
104 **Higuchi M**, Kazazian HH, Jr, Kasch L, *et al.* Molecular characterization of severe hemophilia A suggests that about half the mutations are not within the coding regions and splice junctions of the factor VIII gene. *Proc Natl Acad Sci U S A* 1991;**88**:7405–9.
105 **Green PM**, Bentley DR, Mibashan RS, *et al.* Molecular pathology of haemophilia B. *EMBO J* 1989;**8**:1067–72.
106 **Koeberl DD**, Bottema CD, Sarkar G, *et al.* Recurrent nonsense mutations at arginine residues cause severe hemophilia B in unrelated hemophiliacs. *Hum Genet* 1990;**84**:387–90.
107 **Green PM**, Montandon AJ, Bentley DR, *et al.* The incidence and distribution of CpG→TpG transitions in the coagulation factor IX gene. A fresh look at CpG mutational hotspots. *Nucleic Acids Res* 1990;**18**:3227–31.
108 **Schwaab R**, Ludwig M, Kochhan L, *et al.* Detection and characterisation of two missense mutations at a cleavage site in the factor VIII light chain. *Thromb Res* 1991;**61**:225–34.
109 **Hamaguchi M**, Matsushita T, Tanimoto M, *et al.* Three distinct point mutations in the factor IX gene of three Japanese CRM+ hemophilia B patients (factor IX BMNagoya 2, factor IX Nagoya 3 and 4). *Thromb Haemost* 1991;**65**:514–20.
110 **Bertina RM**, van dL, I, Mannucci PM, *et al.* Mutations in hemophilia Bm occur at the Arg180–Val activation site or in the catalytic domain of factor IX. *J Biol Chem* 1990;**265**:10876–83.
111 **Noyes CM**, Griffith MJ, Roberts HR, *et al.* Identification of the molecular defect in factor IX Chapel Hill: substitution of histidine for arginine at position 145. *Proc Natl Acad Sci U S A* 1983;**80**:4200–2.
112 **Toomey JR**, Stafford D, Smith K. Factor IX Albuquerque (arginine 145 to cysteine) is cleaved slowly by factor XIa and has reduced coagulant activity. *Blood* 1988;**72**:312a.
113 **Monroe DM**, McCord DM, Huang MN, *et al.* Functional consequences of an arginine 180 to glutamine mutation in factor IX Hilo. *Blood* 1989;**73**:1540–4.
114 **Hamer RJ**, Koedam JA, Beeser-Visser NH, *et al.* Factor VIII binds to von Willebrand factor via its Mr-80,000 light chain. *Eur J Biochem* 1987;**166**:37–43.
115 **Shima M**, Scandella D, Yoshioka A, *et al.* A factor VIII neutralizing monoclonal antibody and a human inhibitor alloantibody recognizing epitopes in the C2 domain inhibit factor VIII binding to von Willebrand factor and to phosphatidylserine. *Thromb Haemost* 1993;**69**:240–6.
116 **Saenko EL**, Scandella D. The acidic region of the factor VIII light chain and the C2 domain together form the high affinity binding site for von willebrand factor. *J Biol Chem* 1997;**272**:18007–14.
117 **Jacquemin M**, Lavend'homme R, Benhida A, *et al.* A novel cause of mild/moderate hemophilia A: mutations scattered in the factor VIII C1 domain reduce factor VIII binding to von Willebrand factor. *Blood* 2000;**96**:958–65.
118 **Leyte A**, Van Schijndel HB, Niehrs C, *et al.* Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. *J Biol Chem* 1991;**266**:740–6.
119 **Higuchi M**, Wong C, Kochhan L, *et al.* Characterization of mutations in the factor VIII gene by direct sequencing of amplified genomic DNA. *Genomics* 1990;**6**:65–71.
120 **Pittman DD**, Wang JH, Kaufman RJ. Identification and functional importance of tyrosine sulfate residues within recombinant factor VIII. *Biochemistry* 1992;**31**:3315–25.
121 **Gilles JG**, Lavend'homme R, Peerlinck K, *et al.* Some factor VIII (FVIII) inhibitors recognise a FVIII epitope(s) that is present only on FVIII-vWF complexes. *Thromb Haemost* 1999;**82**:40–5.
122 **Nishino M**, Girma JP, Rothschild C, *et al.* New variant of von Willebrand disease with defective binding to factor VIII. *Blood* 1989;**74**:1591–9.
123 **Mazurier C**. von Willebrand disease masquerading as haemophilia A. *Thromb Haemost* 1992;**67**:391–6.
124 **Roelse JC**, De Laaf RT, Timmermans SM, *et al.* Intracellular accumulation of factor VIII induced by missense mutations Arg593→Cys and Asn618→Ser explains cross-reacting material-reduced haemophilia A. *Br J Haematol* 2000;**108**:241–6.
125 **Frazier D**, Smith KJ, Cheung WF, *et al.* Mapping of monoclonal antibodies to human factor IX. *Blood* 1989;**74**:971–7.
126 **Ludwig M**, Sabharwal AK, Brackmann HH, *et al.* Hemophilia B caused by five different nondeletion mutations in the protease domain of factor IX. *Blood* 1992;**79**:1225–32.
127 **Mathur A**, Zhong D, Sabharwal AK, *et al.* Interaction of factor IXa with factor VIIIa. Effects of protease domain Ca^{2+} binding site, proteolysis in the autolysis loop, phospholipid, and factor X. *J Biol Chem* 1997;**272**:23418–26.
128 **Hamaguchi N**, Bajaj SP, Smith KJ, *et al.* The role of amino-terminal residues of the heavy chain of factor IXa in the binding of its cofactor, factor VIIIa. *Blood* 1994;**84**:1837–42.
129 **Oldenburg J**, Quenzel EM, Harbrecht U, *et al.* Missense mutations at ALA-10 in the factor IX propeptide: an insignificant variant in normal life but a decisive cause of bleeding during oral anticoagulant therapy. *Br J Haematol* 1997;**98**:240–4.
130 **Jorgensen MJ**, Cantor AB, Furie BC, *et al.* Recognition site directing vitamin K-dependent gamma-carboxylation resides on the propeptide of factor IX. *Cell* 1987;**48**:185–91.
131 **Rabiet MJ**, Jorgensen MJ, Furie B, *et al.* Effect of propeptide mutations on post-translational processing of factor IX. Evidence that beta-hydroxylation and gamma-carboxylation are independent events. *J Biol Chem* 1987;**262**:14895–8.
132 **Stenflo J**, Ferlund P, Egan W, *et al.* Vitamin K dependent modifications of glutamic acid residues in prothrombin. *Proc Natl Acad Sci U S A* 1974;**71**:2730–3.
133 **Lingenfelter SE**, Berkner KL. Isolation of the human gamma-carboxylase and a gamma-carboxylase-associated protein from factor IX-expressing mammalian cells. *Biochemistry* 1996;**35**:8234–43.
134 **Willingham AK**, Matschiner JT. Changes in phylloquinone epoxidase activity related to prothrombin synthesis and microsomal clotting activity in the rat. *Biochem J* 1974;**140**:435–41.
135 **Chu K**, Wu SM, Stanley T, *et al.* A mutation in the propeptide of factor IX leads to warfarin sensitivity by a novel mechanism. *J Clin Invest* 1996;**98**:1619–25.
136 **Crossley M**, Ludwig M, Stowell KM, *et al.* Recovery from hemophilia B Leyden: an androgen-responsive element in the factor IX promoter. *Science* 1992;**257**:377–9.
137 **Briet E**, Bertina RM, van Tilburg NH, *et al.* Hemophilia B Leyden: a sex-linked hereditary disorder that improves after puberty. *N Engl J Med* 1982;**306**:788–90.
138 **Crossley M**, Brownlee GG. Disruption of a C/EBP binding site in the factor IX promoter is associated with haemophilia B. *Nature* 1990;**345**:444–6.
139 **Morgan GE**, Rowley G, Green PM, *et al.* Further evidence for the importance of an androgen response element in the factor IX promoter. *Br J Haematol* 1997;**98**:79–85.
140 **Higuchi M**, Antonarakis SE, Kasch L, *et al.* Molecular characterization of mild-to-moderate hemophilia A: detection of the mutation in 25 of 29 patients by denaturing gradient gel electrophoresis. *Proc Natl Acad Sci U S A* 1991;**88**:8307–11.
141 **Goodeve AC**, Williams I, Bray GL, *et al.* Relationship between factor VIII mutation type and inhibitor development in a cohort of previously untreated patients treated with recombinant factor VIII (Recombinate). Recombinate PUP study group. *Thromb Haemost* 2000;**83**:844–8.
142 **Vidal F**, Elisenda F, Carme A, *et al.* Rapid hemophilia A molecular diagnosis by a simple DNA sequencing procedure: identification of 14 novel mutations. *Thromb Haemost* 2001;**85**:580–3.
143 **Naylor JA**, Green PM, Montandon AJ, *et al.* Detection of three novel mutations in two haemophilia A patients by rapid screening of whole essential region of factor VIII gene. *Lancet* 1991;**337**:635–9.
144 **Montandon AJ**, Green PM, Giannelli F, *et al.* Direct detection of point mutations by mismatch analysis: application to haemophilia B. *Nucleic Acids Res* 1989;**17**:3347–58.
145 **Green PM**, Montandon AJ, Ljung R, *et al.* Haemophilia B mutations in a complete Swedish population sample: a test of new strategy for the genetic counselling of diseases with high mutational heterogeneity. *Br J Haematol* 1991;**78**:390–7.
146 **Saad S**, Rowley G, Tagliavacca L, *et al.* First report on UK database of haemophilia B mutations and pedigrees. UK haemophilia centres. *Thromb Haemost* 1994;**71**:563–70.
147 **Waseem NH**, Bagnall R, Green PM, *et al.* Start of UK confidential haemophilia A database: analysis of 142 patients by solid phase fluorescent chemical cleavage of mismatch. Haemophilia centres. *Thromb Haemost* 1999;**81**:900–5.
148 **Rowley G**, Saad S, Giannelli F, *et al.* Ultrarapid mutation detection by multiplex, solid-phase chemical cleavage. *Genomics* 1995;**30**:574–82.
149 **Costa JM**, Ernault P, Vidaud D, *et al.* Fast and efficient mutation detection method using multiplex PCR and cycle sequencing—application to haemophilia B. *Thromb Haemost* 2000;**83**:244–7.
150 **Giannelli F**, Green PM, Naylor JA. A genetic view on the etiology of the inhibitor complication. *Blood* 1996;**87**:2612.
151 **Tuddenham EG**, McVey JH. The genetic basis of inhibitor development in haemophilia A. *Haemophilia* 1998;**4**:543–5.
152 **Schwaab R**, Brackmann HH, Meyer C, *et al.* Haemophilia A: mutation type determines risk of inhibitor formation. *Thromb Haemost* 1995;**74**:1402–6.
153 **Tuddenham EG**. Molecular biological aspects of inhibitor development. *Vox Sang* 1999;**77**(suppl 1):13–16.
154 **Vianello F**, Radossi P, Tison T, *et al.* Prevalence of anti-FVIII antibodies in severe haemophilia A patients with inversion of intron 22. *Br J Haematol* 1997;**97**:807–9.
155 **Hay CR**, Ollier W, Pepper L, *et al.* HLA class II profile: a weak determinant of factor VIII inhibitor development in severe haemophilia A. UKHCDO inhibitor working party. *Thromb Haemost* 1997;**77**:234–7.
156 **Oldenburg J**, Picard JK, Schwaab R, *et al.* HLA genotype of patients with severe haemophilia A due to intron 22 inversion with and without inhibitors of factor VIII. *Thromb Haemost* 1997;**77**:238–42.
157 **Oldenburg J**, Schroder J, Schmitt C, *et al.* Small deletion/insertion mutations within poly-A runs of the factor VIII gene mitigate the severe haemophilia A phenotype. *Thromb Haemost* 1998;**79**:452–3.
158 **Young M**, Inaba H, Hoyer LW, *et al.* 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**:565–73.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

159 **Lenting PJ**, Neels JG, van den Berg BM, *et al*. The light chain of factor VIII comprises a binding site for low density lipoprotein receptor-related protein. *J Biol Chem* 1999;**274**:23734–9.
160 **Saenko EL**, Yakhyaev AV, Mikhailenko I, *et al*. Role of the low density lipoprotein-related protein receptor in mediation of factor VIII catabolism. *J Biol Chem* 1999;**274**:37685–92.
161 **Schwarz HP**, Lenting PJ, Binder B, *et al*. Involvement of low-density lipoprotein receptor-related protein (LRP) in the clearance of factor VIII in von Willebrand factor-deficient mice. *Blood* 2000;**95**:1703–8.
162 **Neels JG**, van den Berg BM, Mertens K, *et al*. Activation of factor IX zymogen results in exposure of a binding site for low-density lipoprotein receptor-related protein. *Blood* 2000;**96**:3459–65.

ECHO

Setting the pace



Please visit the Molecular Pathology website [www.jclinpath.com] for link to this full article.

The possibility of gene therapy to treat cardiac chronotropic incompetence took a step forward with the successful, if transient, enhancement of heart rate in a large animal model by human gene transfer.

Having already established the principle in a mouse model, with transfer of constructs encoding the β_2 human adrenergic receptor resulting in enhanced heart rate, Edelberg and colleagues repeated the process in a Yorkshire pig model — chosen for its similarity to the human cardiovascular system. Plasmids encoding either human β_2 adrenergic receptor (pBR322-β actin promoter-β_2 adrenergic receptor) or control constructs (pBR322-β actin promoter-β galactosidase) and a co-injection vector (plasmid construct encoding humanised green fluorescent protein with a cytomegalovirus promoter) were injected by cardiac catheterisation into the right atrial myocardium (six pigs for β_2 adrenergic receptor and five for controls) in a randomised double blind procedure.

Two days later the average heart rate before injection (108 (SD 16) beats/min) was significantly increased in pigs receiving the β_2 adrenergic receptor construct (163 (33) beats/min; $p<0.05$ compared with preinjection) compared with those receiving control constructs (127(25) beats/min; $p>0.3$ compared with preinjection). The increase before and after injection was almost a 50% increase in heart rate ($p<0.01$ compared with that for the controls). After sacrifice sections of the injection sites of the excised hearts were stained with immunofluorescent stain specific for the human β_2 adrenergic receptor. Dual fluorescence microscopy showed specific staining only in those hearts injected with the β_2 adrenergic receptor construct, indicating expression of the encoded genes.

▲ *Heart* 2001;**86**:559–562.

Mol Path: first published as 10.1136/mp.55.1.1 on 1 February 2002. Downloaded from http://mp.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

REVIEW (corrected version)

Haemophilia A and haemophilia B: molecular insights

D J Bowen

J Clin Pathol: Mol Pathol 2002;**55**:127–144

This review focuses on selected areas that should interest both the scientist and the clinician alike: polymorphisms within the factor VIII and factor IX genes, their linkage, and their ethnic variation; a general assessment of mutations within both genes and a detailed inspection of the molecular pathology of certain mutations to illustrate the diverse cause–effect relations that exist; a summary of current knowledge on molecular aspects of inhibitor production; and an introduction to the new areas of factor VIII and factor IX catabolism. An appendix defining various terms encountered in the molecular genetics of the haemophilias is included, together with an appendix providing accession numbers and locus identification links for accessing gene and sequence information in the international nucleic acid databases.

Correction

Haemophilia A and haemophilia B: molecular insights. Bowen DJ. *Mol Pathol* 2002;**55**:1–18.

Owing to an error in the electronic processing of this paper, figures 4 and 11 were incorrect. Because this was entirely the fault of the journal and not the author this paper is being republished in full. When citing this paper please use the new page numbers.

Correspondence to:
D J Bowen, Department of Haematology, University of Wales College of Medicine, Heath Park, Cardiff CF3 0DQ, South Wales, UK;
bowendj1@cf.ac.uk

Accepted for publication
10 May 2001

The molecular basis of the hereditary haemophilias has been unravelled to a remarkable degree over the past decade. The early studies, which glimpsed the possibility of mutational heterogeneity, have been verified by the almost breath taking diversity of mutations that have since been characterised in the coagulation factor VIII gene in haemophilia A and the coagulation factor IX gene in haemophilia B. Although mutational heterogeneity abounds within these two genes, natural polymorphic heterogeneity does not. These apposite features have made the collection of informative genetic data for familial investigations a difficult task. However, modern techniques of molecular biology have decreased the enormity of the task of mutation analysis in these two genes and have decreased the workload for the analysis of their polymorphisms. This review focuses on polymorphisms and mutations in the genes for coagulation factor VIII and factor IX, molecular biological aspects of inhibitor production, and the new areas of factor VIII and factor IX catabolism. The molecular pathology of certain mutations is dealt with to exemplify the interesting and revealing cause–effect relations that exist within the haemophilias.

> "An almost breath taking diversity of mutations have since been characterised in the coagulation factor VIII gene in haemophilia A and the coagulation factor IX gene in haemophilia B"

BACKGROUND: FACTOR VIII AND FACTOR IX IN COAGULATION

These two proteins circulate as inactive precursors that are activated at the time of haemostatic challenge, via the intrinsic or extrinsic pathways of the coagulation cascade (fig 1). Factor VIII is a protein cofactor with no enzyme activity per se; factor IX is a serine protease with an absolute requirement for factor VIII as cofactor. Upon activation, and in the presence of calcium ions and phospholipid surfaces, factor VIII and factor IX form an active complex, the tenase complex, which activates factor X (fig 1). Subsequent stages of the cascade then proceed, culminating in the deposition of fibrin, the structural polymer of the blood clot.

A deficiency or dysfunction of either factor VIII or factor IX compromises the activation of factor X, so that the ensuing steps of the coagulation cascade are also compromised and fibrin deposition is either inefficient or non-existent. Thus, the fundamental biochemical lesion underlying the haemophilias is an insufficiency of the activity of the tenase complex, brought about either by a deficiency of coagulation factor VIII cofactor activity (haemophilia A) or coagulation factor IX enzyme activity (haemophilia B). Thus, it is not surprising that the two disorders are clinically similar because they both arise from perturbation of the same essential step in the process of fibrin generation.

THE FACTOR VIII AND FACTOR IX GENES AND PROTEINS

Both genes are located on the X chromosome: the factor VIII gene is located towards the end of the long arm at Xq28[1]; the factor IX gene is also located on the long arm, more towards the centromere at Xq27.[2] There is a considerable genetic distance (35 cM) between the two genes[3 4]; they are by no means in tight linkage with each other. Thus, haemophilia A and B are X linked disorders; they are recessive, carried by females (karyotype 46:XX), and present in males (karyotype 46:XY). Rare cases of female haemophilia are known[5 6] and these may arise through the presence of two defective factor VIII or factor IX genes, or through non-random X chromosome inactivation. The latter has often been offered as a primary explanation for haemophilia symptoms in known carriers; however, recent data suggest that there is an absence of correlation between X chromosome inactivation and plasma concentrations of factor VIII in carriers of haemophilia A and B.[7]

Abbreviations: ARE, androgen response element; C/EBP, CCAAT/enhancer binding protein; EGF, epidermal growth factor; LD, long distance; LRP, lipoprotein receptor related protein; MHC, major histocompatibility complex; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; SNP, single nucleotide polymorphism; VNTR, variable number tandem repeat sequence

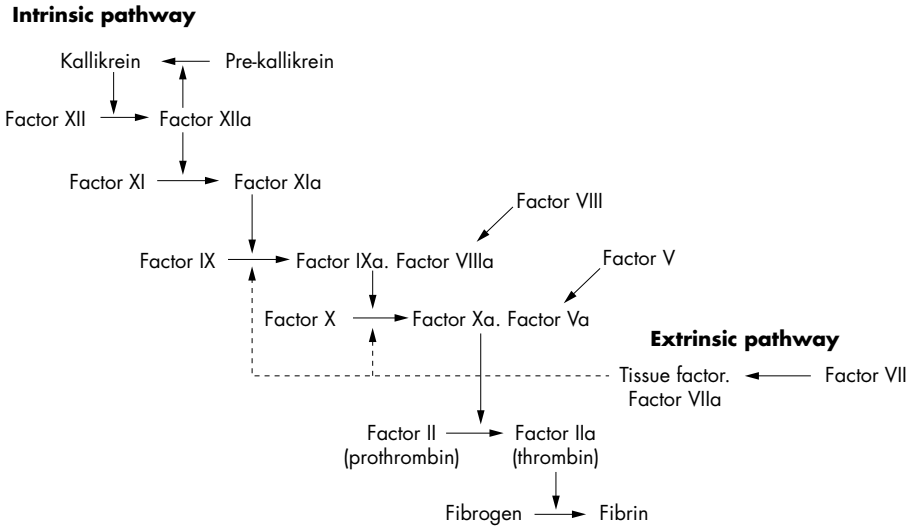


Figure 1 Schematic showing the intrinsic and extrinsic pathways of the coagulation cascade leading to fibrin formation. A deficiency or dysfunction of coagulation factor VIII or factor IX compromises the activation of factor X, the ensuing reactions are inefficient and haemophilia results.

Factor VIII

The factor VIII gene is extremely large (~180 kb) and structurally complex (26 exons) (fig 2A; table 1).[1] Intron 22 of the gene contains a CpG island, which acts as a bidirectional promoter for two further genes, F8A[8] and F8B[9] (fig 2B). The CpG island and F8A are contained within a stretch of DNA of approximately 9.5 kb, which is repeated at least twice on the X chromosome, further towards the telomere and extragenic to the factor VIII gene[8 10] (fig 2C). These homologues are known as int22h-1 (intragenic) and int22h-2 and int22h-3 (extragenic).[10]

F8A is intronless and spans less than 2 kb within intron 22.[8] It is transcribed in the direction opposite to factor VIII.[8] It is transcribed abundantly in a wide variety of cell types[8] and is conserved in the mouse, which implies it has some function.[11]

F8B comprises a 5′ exon located within intron 22 and then exons 23–26 of the factor VIII gene.[10] The 5′ exon potentially encodes eight amino acids and the factor VIII reading frame is maintained for exons 23–26.[10] The result is a protein that is considerably shorter than factor VIII and that includes the phospholipid binding domain.

The factor VIII gene transcript is approximately 9010 bases in length and comprises a short 5′ untranslated region (150 bases), an open reading frame plus stop codon (7056 bases), and an unusually long 3′ untranslated region (1806 bases)[12] (fig 3A). The open reading frame encodes a signal peptide of 19 amino acids, which directs the passage of factor VIII through the cell, and a mature protein of 2332 amino acids, which is the circulating inactive pro-cofactor (fig 3B).[12] This contains a distinct domain structure based on homologies within the protein.[13]

Factor VIII is activated by proteolysis catalysed by thrombin (fig 3C) (reviewed in Tuddenham and Cooper[14]). The principal activation cleavage sites are on the C-terminal side of arginine residues 372, 740, and 1689.[15] Activated factor X also cleaves at these sites. The activation cleavage sites flank the B domain, which is released from factor VIII on activation, leaving a heterotrimer comprising an N-terminal heavy chain and a

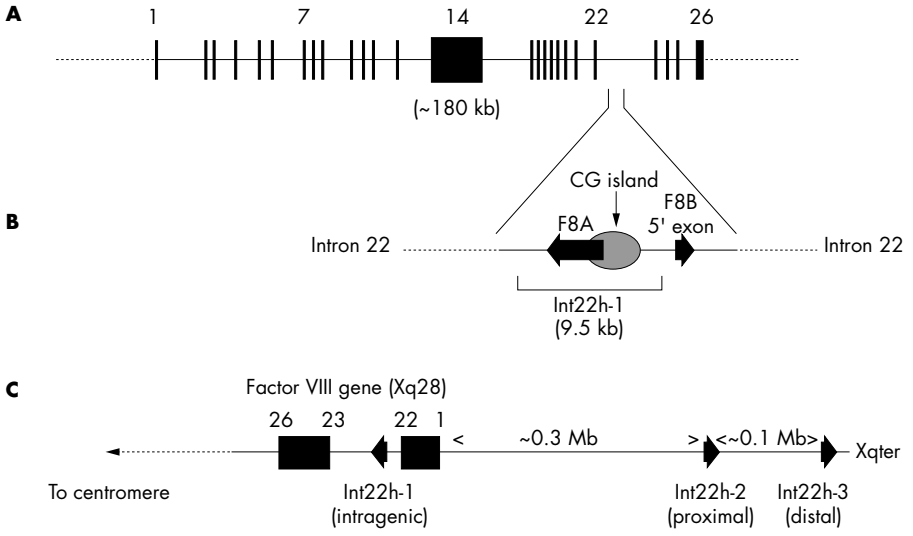


Figure 2 (A) Genomic organisation of the human factor VIII gene. Exon and intron sizes are given in table 1. (B) Enlargement of a portion of intron 22 of the factor VIII gene to show the relative locations and orientations of the CG island, the intragenic gene F8A, the 5′ exon of the putative gene F8B, and the extent of the int22h-1 homologue, which is duplicated towards the telomere of Xq. (C) Orientation of the factor VIII gene within Xq28 and location and orientation of int22h homologues.

Table 1 Exon and intron sizes for the human factor VIII gene[1]

Exon	Length (bp)	Intron	Length (kb)
1	313	1	22.9
2	122	2	2.6
3	123	3	3.9
4	213	4	5.4
5	69	5	2.4
6	117	6	14.2
7	222	7	2.6
8	262	8	0.3
9	172	9	4.8
10	94	10	3.8
11	215	11	2.8
12	151	12	6.3
13	210	13	16.0
14	3106	14	22.7
15	154	15	1.3
16	213	16	0.3
17	229	17	0.2
18	183	18	1.8
19	117	19	0.6
20	72	20	1.6
21	86	21	3.4
22	156	22	32.4
23	145	23	1.4
24	149	24	1.0
25	177	25	22.4
26	1958		

C-terminal light chain; these are held together by Ca^{2+}. In the circulation, factor VIII is carried and protected by von Willebrand factor.[16 17] Proteolytic cleavage of factor VIII at the time of activation simultaneously releases it from its complex with von Willebrand factor.[18 19]

"In the circulation, factor VIII is carried and protected by von Willebrand factor"

Factor IX

The factor IX gene is considerably smaller and structurally simpler than that for factor VIII. It is approximately 34 kb in length and contains only eight exons, the largest of which is only 1935 bp (fig 4A; table 2).[2] The transcript is 2803 bases in length and comprises a short 5′ untranslated region (29 bases), an open reading frame plus stop codon (1383 bases), and a 3′ untranslated region (1390 bases) (fig 4B).[20] The open reading frame encodes a pre-pro-protein in which the pre-sequence (or signal sequence) directs factor IX for secretion, the pro-sequence provides a binding domain for a vitamin K dependent carboxylase, which carboxylates certain glutamic acid residues in the adjacent GLA domain, and the remainder represents the factor IX zymogen, which enters the circulation after removal of the pre- and pro-sequences. Domains within the zymogen are identified according to structure or function.

Activation of factor IX involves cleavage of two peptide bonds, one on the C-terminal side of arginine 145 (the α-cleavage) the other on the C-terminal side of arginine 180 (the β-cleavage) (fig 4C).[21] These cleavages are brought about by activated factor XI generated through the intrinsic pathway,[22] or via the tissue factor/activated factor VII complex of the extrinsic pathway.[23 24] The relative importance of these two pathways to factor IX activation is indicated by the fact that severe factor XI deficiency is typically associated with a mild bleeding disorder, whereas severe factor VII deficiency is characterised by a severe bleeding disorder. The activation cleavages generate an N-terminal light chain and a C-terminal heavy chain, held together by a disulphide bond between cysteine residues 132 and 279 (fig 4D).[21]

The above overview is necessarily brief. Its purpose is to relate salient features that provide the context and framework for the molecular information that follows. More in depth coverage can be obtained from several excellent reviews.[14 25–28]

POLYMORPHISMS

The factor VIII and factor IX genes both contain two types of polymorphism: single nucleotide polymorphisms (SNPs) and length polymorphisms, also known as variable number

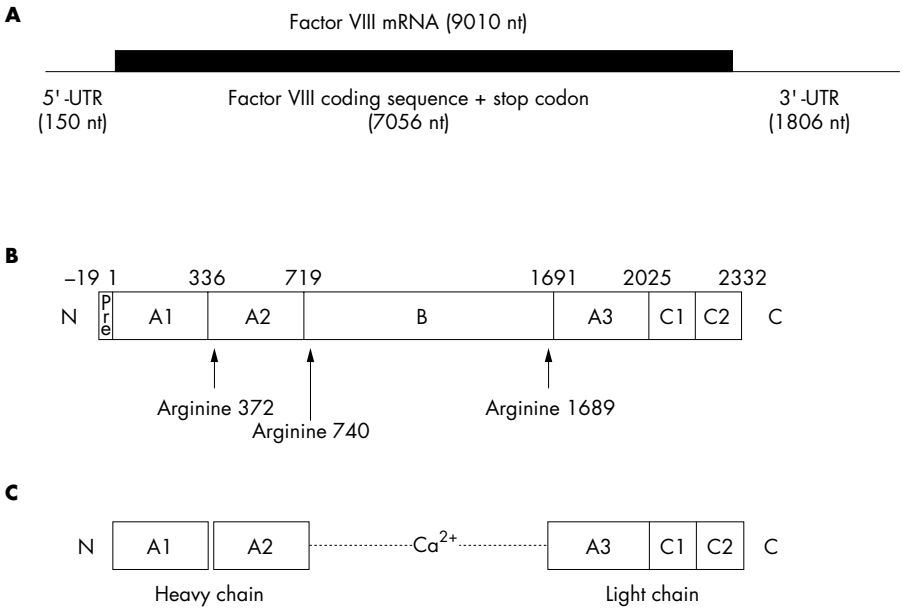


Figure 3 (A) Factor VIII mRNA showing the extent and location of the open reading frame. (B) The newly synthesised factor VIII protein molecule comprising a pre-sequence of 19 amino acids and a mature peptide of 2332 amino acids (total length, 2351 amino acids). A1–3, B, C1, and C2 represent domains assigned according to homologies within the protein (the boundaries are somewhat arbitrary). The arginine residues signalling the sites for proteolytic activation are arrowed. (C) Activated factor VIII comprising a heterotrimer in which the dimeric N-terminal heavy chain is held together with the monomeric C-terminal light chain by a metal ion bridge (Ca^{2+}).

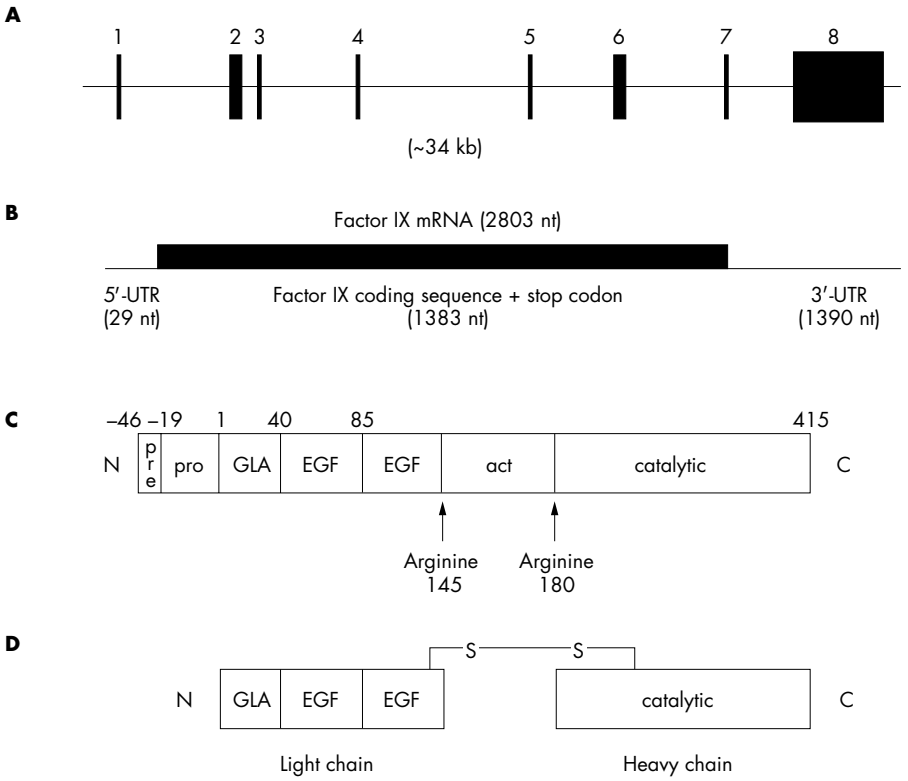


Figure 4 (A) Genomic organisation of the human factor IX gene. Exon and intron sizes are given in table 2. (B) Factor IX mRNA showing the relative size and location of the open reading frame. (C) The newly synthesised factor IX protein molecule comprising a pre- and pro-sequence (27 and 19 amino acids, respectively) and a mature peptide of 415 amino acids (total length, 461 amino acids). (D) Activated factor IX comprising an N-terminal light chain and a C-terminal heavy chain held together by a disulphide bridge between cysteine resides 132 and 279. GLA, "GLA" domain, in which 12 glutamic acid residues undergo post-translational γ-carboxylation by a vitamin K dependent carboxylase; EGF, epidermal growth factor-like domain; act, activation peptide released after proteolytic activation at arginine 145 and arginine 180; catalytic, the serine protease domain.

tandem repeat sequences (VNTRs) or microsatellites. Polymorphisms have a scientific interest of their own and are useful in areas as diverse as forensic science[29 30] and the study of human evolution.[31 32] They have clinical relevance in the context of hereditary disorders in that they can be used to track a defective (or normal) gene through an affected family. Such linkage studies have permitted carrier status investigation and prenatal diagnosis in haemophilia A and haemophilia B.[33 34]

There is an apparent paucity of polymorphisms in both the factor VIII and factor IX genes. In the recent presentation of the sequence of the human genome,[35] two haploid human genomes chosen randomly differed by 1 base/1250 bases on average; therefore, the factor VIII and factor IX genes would be expected to contain approximately 144 and 27 SNPs, respectively. However, these numbers are considerably greater than the number of polymorphisms detected so far in these genes, although this could reflect a problem of detection rather than a genuine paucity. Certainly, many more candidate polymorphisms have been identified as a result of the human genome sequencing programme[35 36] than the recognised polymorphisms described below.

Table 2 Exon and intron sizes for the human factor IX gene[2]

Exon	Length (bp)	Intron	Length (bp)
1	117	1	6206
2	164	2	188
3	25	3	3689
4	114	4	7163
5	129	5	2565
6	203	6	9473
7	115	7	668
8	1935		

"There is an apparent paucity of polymorphisms in both the factor VIII and factor IX genes"

FACTOR VIII GENE POLYMORPHISMS

The factor VIII gene contains several SNPs, many of which fall into the subcategory of restriction fragment length polymorphisms (RFLPs) (fig 5A). In addition, two microsatellites have been reported within the gene, one in intron 13,[38] the other in intron 22 (fig 5A).[44]

The intron 13 microsatellite comprises the simple repeat (CA)n,[38] whereas the intron 22 comprises (CA)n(CT)n.[44] Differences in the value of "n" at such repeats have been attributed to DNA polymerase slippage and appear to drift towards an increase in repeat length.[57] Because such events could occur between generations within a family, microsatellite data should be interpreted with appropriate caution in a linkage study.[58] Preferably, verification of microsatellite data should be sought from at least one other intragenic locus.

The XbaI and MspI RFLPs in intron 22 of the factor VIII gene (designated XbaI A and MspI A) occur within the int22h-1 sequence.[10 42 43] The corresponding positions in the extragenic homologues int22h-2/3 are also polymorphic for XbaI[42] and MspI[43] (designated XbaI B and C, and MspI B and C, respectively). Caution must therefore be exercised in the use of these two highly informative loci.

The allele frequencies of some of the polymorphisms have been investigated in various populations and differences are

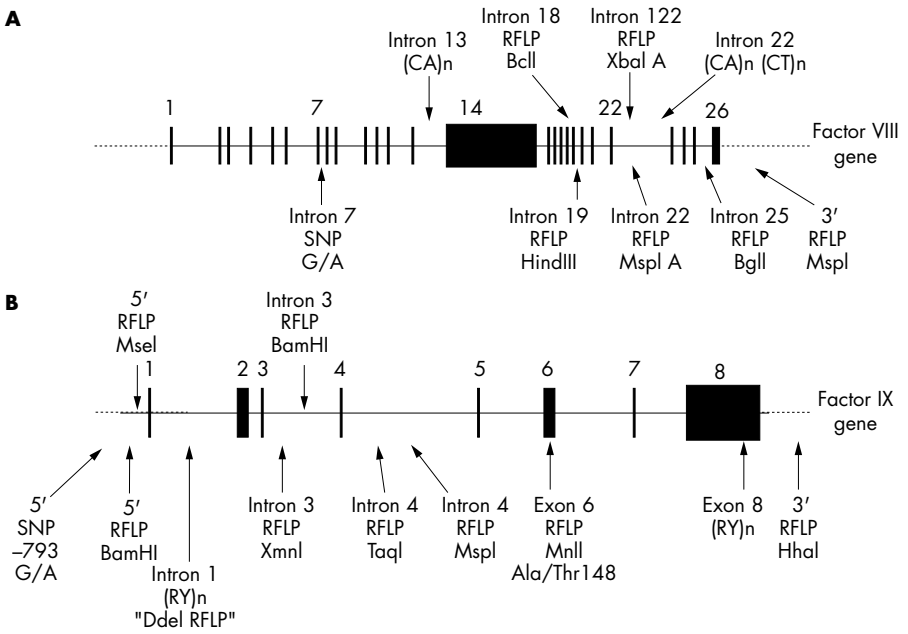


Figure 5 Some of the known polymorphisms in the human genes for (A) factor VIII and (B) factor IX. Source references: factor VIII gene intron 7 G/A,[37] intron 13 (CA)n,[38] intron 18 BclI,[39] intron 19 HindIII,[40] intron 22 XbaI A,[41 42] intron 22 MspI A,[43] intron 22 (CA)n,[44] intron 25 BglI,[45] 3′ MspI,[46] factor IX gene 5′ –793 G/A,[47] 5′ BamHI,[48] 5′ MseI,[49] intron 1 DdeI,[50] intron 3 XmnI,[50] intron 3 BamHI,[51] intron 4 TaqI,[52] intron 4 MspI,[53] exon 6 MnlI,[54] exon 8 (RY)n,[55] and 3′ HhaI.[56]

seen (table 3). Three observations are of particular interest: first, the [+] alleles for the BclI and HindIII RFLPs have very different frequencies in American blacks compared with other ethnic groups, affording a small glimpse into human evolutionary divergence. Second, although whites and Chinese have similar allele frequencies at the BclI, HindIII, and XbaI A loci, they differ noticeably at the BglI locus. The BglI RFLP is the result of a transition that is not likely to recur independently.[62] Thus, it is possible that its presence in whites but not in Chinese might reflect its emergence in the former after the two ethnic groups diverged. However, the presence of the BglI RFLP in Japanese and American blacks suggests that this explanation is too simplistic. Third, in American blacks the allele frequencies for the BclI and HindIII loci are the reverse of what is seen in other ethnic groups; however, for the BglI locus, this is not the case. This complicates the interpretation of the ancestral factor VIII gene haplotype(s) from which the haplotypes existing in present day populations must have evolved.

The differences in allele frequencies between populations have implications for linkage studies in haemophilia A: the loci of greatest diagnostic potential may differ according to the ethnicity of the family to be investigated and this has to be borne in mind in assessing the investigative strategy.

Linkage disequilibrium has not been investigated for all loci, although it has been shown for some. There is strong linkage disequilibrium between the intron 18 BclI, intron 7 G/A, intron 19 HindIII, and intron 25 BglI RFLPs.[59 63 64] The linkage disequilibrium in these cases is strong but not complete; therefore, additional informativity might be obtained from their inclusion in a linkage study. The intron 18 BclI and intron 22 XbaI A loci also show linkage disequilibrium[41]; however, this is less pronounced and their combined analysis provides much more informativity than either used on its own. For example, in whites, the informativity for combined use *v* individual use of XbaI A is approximately 65% *v* 49%,[41] in Chinese it is 52% *v* 49%,[61] and in Japanese it is 79% *v* 48%.[60]

Table 3 Allele frequencies for four factor VIII gene polymorphisms frequently used in linkage studies in haemophilia A

			Frequency			
Location	Name	Alleles	Whites	Japanese	Chinese	American blacks
Intron 18	BclI	+	0.76	0.70	0.82	0.20
		–	0.24	0.30	0.18	0.80
Intron 19	HindIII	+	0.25		0.24	0.78
		–	0.75		0.76	0.22
Intron 22	XbaI A	+	0.58	0.59	0.57	
		–	0.42	0.41	0.43	
Intron 25	BglI	+	0.90	0.91	1	0.74
		–	0.10	0.09	0	0.26

Source references, white: BclI,[59] HindII,[59] XbaI A,[42] BglI,[45]; all Japanese data[60]; Chinese: BclI,[61] HindIII,[59] XbaI,[61]; American black: BclI,[59] HindII,[59] and BglI.[45]

Table 4 Allele frequencies for several factor IX gene polymorphisms in different ethnic groups

Location	Name	Alleles	Frequency			
			Whites	Japanese	Chinese	American blacks
5′	BamHI	+	0.02		0	0.36
		–	0.98		1	0.64
Intron 3	XmnI	+	0.29	0	0.01	0.12
		–	0.71	1	0.99	0.88
Intron 3	BamHI	+	0.06		0.02	0.13
		–	0.94		0.98	0.87
Intron 4	TaqI	+	0.35	0	0.01	0.14
		–	0.65	1	0.99	0.86
Intron 4	MspI	+	0.80	1	0	0.61
		–	0.20	0	1	0.39
Exon 6	MnlI	+	0.33		0.03	0.11
		–	0.67		0.97	0.89
3′	HhaI	+	0.61	0.24	0.17	0.57
		–	0.39	0.76	0.83	0.43

Source references: white data,[26] Japanese data,[69 70] Chinese data,[34 71 72] American black data.[34 72–74]

XbaI A and MspI A are located within 737 bases of each other.[43] Despite their close proximity, the RFLPs are not in complete linkage disequilibrium. Haplotype analysis from a white population predicts an informativity of approximately 60% in linkage studies using both, compared with 47% using either on its own.[43]

Linkage disequilibrium, like allele frequencies, differs between ethnic groups and this should also be taken into account in assessing the linkage analysis strategy within a family.[34] Linkage disequilibrium and low rates of heterozygosity decrease the overall potential informativity of the polymorphisms in the factor VIII gene such that, in practice, about 10–15% of linkage studies are non-informative. The addition of further polymorphic loci to the small array already available would therefore be valuable.

FACTOR IX GENE POLYMORPHISMS

The human factor IX gene contains several SNPs, most of which fortuitously fall into the subcategory of RFLPs (fig 5B). The MnlI polymorphism[54] is located in exon 6 and brings about the amino acid substitution threonine/alanine at position 148 (known as the Malmö polymorphism[65]). Therefore, the factor IX protein itself is polymorphic and the Malmö polymorphism can be genotyped at the protein level using antibodies specific for each variant.[66 67] There are two microsatellite loci, one in intron 1[68] (the "DdeI" polymorphism[50]) and one in the 3′ untranslated region of exon 8[55] (fig 5B). These are of the type (RY)n (R, purine; Y, pyrimidine).

As seen for polymorphisms in the factor VIII gene, factor IX gene polymorphisms show ethnic variation, which in some instances is quite pronounced (table 4). For example, the allele frequencies of the BamHI RFLPs in the 5′ region and in intron 3 are reasonable in the American black population, whereas these loci are essentially non-polymorphic in whites, Chinese, and Japanese (table 4). In contrast, the MnlI RFLP in exon 6 is common in whites and American blacks, but rare in Chinese (table 4). These differences in allele frequencies between racial groups impact upon linkage studies and influence which loci should be included for familial investigation in haemophilia B.[34]

The (RY)n microsatellites have been investigated in considerable depth.[55 68 73] The 3′ (exon 8) locus has the consensus $(GT)_a(ATGCGT)_4AG(AC)_bGCAT(AC)_3(AT)_2$ and four alleles have been demonstrated (I–IV) in which the values of "a" and "b" vary.[55] The intron 1 (RY)n motif comprises two units, A and B, with consensus sequence $GT(AC)_3(AT)_3GT(AT)_I$ (I = 4 for the A unit and 5 for the B unit).[68] The polymorphism is caused by variability in the number of A and B units present. Both microsatellites show variation between racial groups.[55 73] Unlike the microsatellites within the factor VIII gene, the variability in the microsatellites of the factor IX gene does not appear to arise as a result of replication slippage. Instead, the patterns of variation observed are compatible with sister chromatid exchange (homologous recombination).[68]

There is strong circumstantial evidence that (RY)n sequences are involved in homologous recombination.[68] In this light, it is interesting to note that the BamHI RFLP which is 5′ to the intron 1 (RY)n is in linkage equilibrium with polymorphisms located between this (RY)n and the exon 8 (RY)n.[68] Similarly, the HhaI RFLP 3′ to the exon 8 (RY)n is in linkage equilibrium with polymorphisms between this (RY)n and the intron 1 (RY)n.[68] In contrast, the polymorphisms between the intron 1 and exon 8 (RY)n are in strong linkage disequilibrium.[26 50 72]

If the (RY)n sequences of the factor IX gene truly are involved in homologous recombination events, this has evident implications for linkage studies (the loss of linkage between polymorphic loci and the defect in the factor IX gene). However, this has not been seen, although the possibility of its occurrence argues for the inclusion of several loci across the factor IX gene in a linkage study.

Linkage disequilibrium has been investigated for some of the factor IX gene polymorphisms. In whites, the TaqI, XmnI, MspI, and MnlI RFLPs show pronounced linkage disequilibrium, such that the use of all four loci would only increase informativity to 55% compared with 45% when using the TaqI site alone.[26 50 72] The DdeI and HhaI polymorphisms, however, show much less disequilibrium,[72] possibly for the reasons discussed above.

In American blacks, the TaqI, MspI, and intron 3 BamHI loci show linkage disequilibrium[51]; however, the 5′ BamHI, intron 1 (RY)n, and XmnI polymorphisms appear to be in equilibrium.[51 72 73] Once again, these observations may be explained by the involvement of (RY)n sequences in homologous recombination.

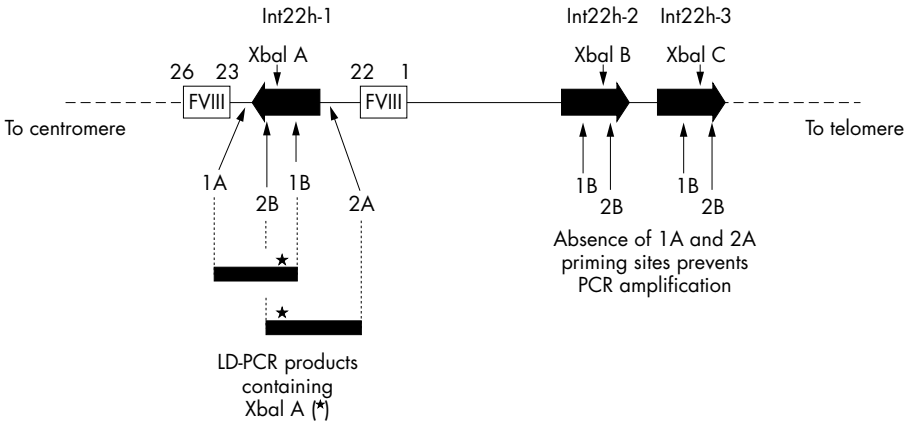


Figure 6 Genotyping the intron 22 XbaI A restriction fragment length polymorphism (RFLP) using long distance PCR (LD-PCR). A primer specific for factor VIII gene sequence (1A or 2A) is used in conjunction with a primer that flanks the XbaI A site and that hybridises within the int22h homologue (1B or 2B, respectively). The resulting LD-PCR product (~ 7 kb) contains XbaI A (asterisk) and can be genotyped by XbaI digestion. The analysis is specific for XbaI A because the sequences to which primers 1A or 2A hybridise do not occur at the extragenic int22h loci. Primers 1B or 2B can hybridise, but in the absence of primers 1A and 2A amplification cannot take place.

"As seen for polymorphisms in the factor VIII gene, factor IX gene polymorphisms show ethnic variation, which in some instances is quite pronounced"

The G/A SNP at −793 is only 104 bases from the MseI RFLP at −698.[49] Loci in such close proximity would be expected to show complete linkage disequilibrium; however, this is not the case. Their combined use provides an informativity of 60% compared with approximately 45% for either marker used alone.[49]

No linkage disequilibrium has been seen between the 3′ HhaI RFLP and other intragenic loci. Taken together with the possible loss of linkage between loci either side of the (RY)n sequences, a theoretical minimum linkage study with high diagnostic potential can be proposed. This would use HhaI 3′ to the exon 8 (RY)n, a second polymorphism 5′ to the intron 1 (RY)n, and a third polymorphism between the (RY)n loci. The second and third loci should be chosen according to racial background.

LINKAGE ANALYSIS IN HAEMOPHILIA A AND B

In essence, the practical approach involves the extraction of genomic DNA from anticoagulated whole blood obtained from relevant family members, amplification of DNA spanning the target polymorphic locus using the polymerase chain reaction (PCR), and then analysis of the PCR product using any of a variety of techniques to distinguish the alleles present[75] (for example, restriction digestion, single strand conformation polymorphism analysis, heteroduplex analysis,[64] PCR product length analysis).

For most of the polymorphic loci in both genes this is a straightforward task and is the mainstream approach in linkage analyses.[34 75] However, for the XbaI A and MspI A RFLPs in intron 22 of the factor VIII gene, the task is complicated by the extragenic homologues (XbaI B and C, and MspI B and C). Until recently, specific amplification of these loci in int22h-1 was not possible. However, the introduction of long distance PCR (LD-PCR), in which stretches of genomic DNA as long as tens of kilobases can be amplified from small amounts of starting material, paved the way for the solution of specific amplification of the int22h-1 loci.[76 77]

The approach is outlined in fig 6. One LD-PCR primer hybridises to all three int22 homologues, whereas a second primer hybridises only to the factor VIII gene sequence outside the int22h-1 region. In the LD-PCR, exponential amplification is only possible for the int22h-1 target as a result of the hybridisation of both primers; int22h-2 and int22h-3 fail to amplify because only one primer can hybridise.

The resulting LD-PCR product can then be genotyped for XbaI A by digestion with this restriction endonuclease,[76 77] or it can be used as the substrate in a second round of PCR that contains primers for amplification of the MspI A site also contained within the LD-PCR product. MspI A can then be genotyped by digestion of the second round product with MspI.[77]

It is clear that PCR has provided a major step forward for the analysis of polymorphisms in the haemophilias. The traditional approach for RFLP analysis, using Southern blot, was labour intensive, time consuming, hazardous, and required relatively large quantities of high molecular weight DNA. In comparison, PCR is rapid, non-labourious, non-hazardous, and requires only small quantities of DNA, which can be degraded to a considerable degree.

Therefore, it is not surprising that the use of PCR is now common to most, if not all, laboratories engaged in genetic testing in the haemophilias. However, the methods used for post-PCR product analysis differ enormously. At one end of the spectrum there are low tech approaches, such as agarose gel electrophoresis followed by ethidium bromide staining, whereas at the other there are the high tech approaches such as the Light Cycler™ or fluorescence based capillary electrophoresis. The choice within laboratories reflects many parameters that need not concern us here; the important point is that post-PCR product analysis is flexible and amenable to the most up to date technologies in molecular biology. Whether "DNA chips"[78–80] (fig 7) may one day become useful for linkage studies in the haemophilias remains to be seen: the evolving movement towards direct mutation detection covered in the next section may ultimately result in linkage analysis being pushed into the sidelines.

THE MOLECULAR BASIS OF THE HAEMOPHILIAS

The mutations causing haemophilia A and B have been localised and characterised in several hundreds of patients. What is immediately evident from the enormous number of mutations that have been elucidated is that the molecular basis of the haemophilias is extremely diverse. This review will focus on some of the principal generalisations that emerge from an inspection of the vast databases of haemophilia mutations and it will consider some specific mutations that illustrate the differing molecular pathologies that exist.

The relation between cause and effect is not always clear, and only in a minority of mutations do we have a good understanding of how haemophilia ensues from the gene defect that

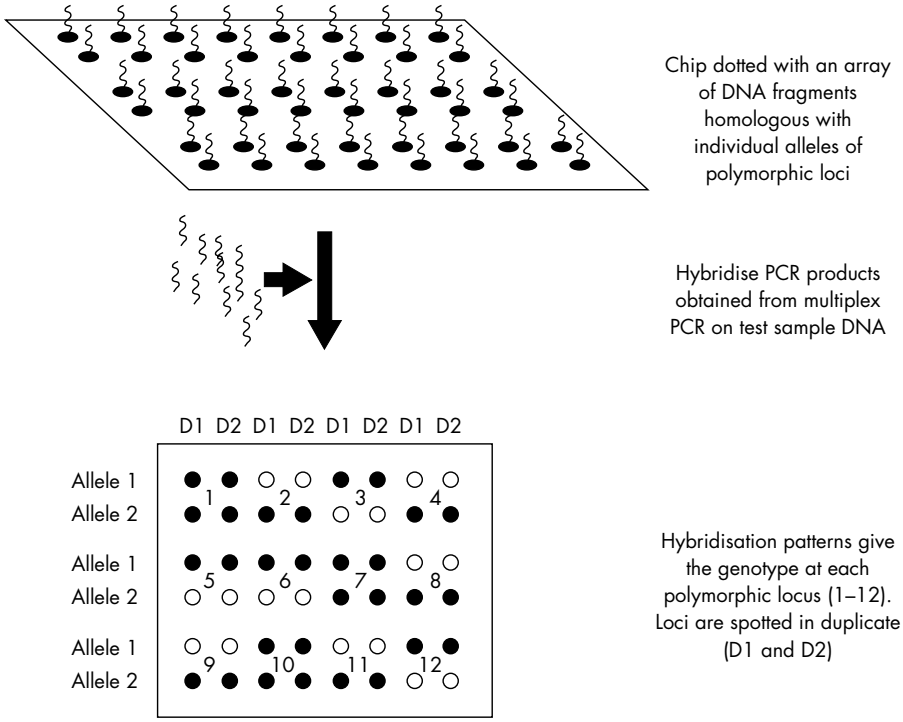


Figure 7 One approach to polymorphism screening using DNA chips. Multiplex PCR could be used to generate locus specific fragments for hybridisation to the chip or, conceivably, fragmented, labelled genomic DNA could be hybridised directly, without amplification.

is present. Unfortunately, this is mostly confined to those mutations whose effect is self evident in that the gene is disrupted, translation is prematurely terminated, activation of the protein is hindered, or mRNA splicing is corrupted, etc. For most mutations, there is no obvious explanation for effect, in which case one is left with candidate explanations such as protein destabilisation, incorrect folding, perturbation of structurally or functionally important regions, and so on. These are often difficult possibilities to prove. However, *x* ray crystallography and advances in software for molecular modelling and prediction of protein structure have permitted these possibilities to be explored for some mutations.[81–84]

OVERVIEW OF MUTATIONS

Mutations that have been characterised in the haemophilias are listed in the following databases on the World Wide Web (these databases are freely accessible).

- Haemophilia A: http://europium.csc.mrc.ac.uk/usr/WWW/WebPages/main.dir/main.htm (HAMSTeRS, The Haemophilia A Mutation, Structure, Test, Resource Site).
- Haemophilia B: http://www.umds.ac.uk/molgen/ (Haemophilia B Mutations Database).
- Both can be accessed via: http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html (Human Gene Mutation Database).

The databases are extensive listings that represent a vast coverage both of published and unpublished findings.[85] From these listings, several generalisations emerge that are common to both haemophilia A and B. The following is a synthesis of those generalisations.

TYPES OF MUTATION

Point mutations, deletions, insertions, and rearrangements/inversions are found in both the factor VIII and factor IX genes. Point mutations (single nucleotide substitutions) are the most common gene defect and are present in approximately 90% of patients. Deletions are the second most common gene defect and are present in approximately 5–10% of patients. Insertions and rearrangements/inversions are quite rare within the haemophilia population, with the exception of an inversion that is prevalent among patients with severe haemophilia A (this inversion, the intron 22 inversion, will be discussed in more detail later).

Point mutations

The point mutations that occur in the haemophilias comprise missense mutations (these change a codon so that a different amino acid is encoded), nonsense point mutations (these change an amino acid codon into a translation stop codon), and mRNA splice site point mutations (these corrupt a true mRNA splice site, or create a novel one).

Missense mutations

The severity of haemophilia arising from missense point mutations depends very much on the nature of the amino acid substitution and its location. Semiconservative amino acid substitutions (the new amino acid is similar to the normal one either with respect to charge, hydrophobicity, polarity, or shape) tend to be associated with less severe disease except when they occur in regions of structural or functional importance. Non-conservative amino acid substitutions (the new amino acid is very dissimilar to the normal amino acid) tend to be associated with moderate to severe disease wherever they occur in the protein. The scope for structural perturbation/destabilisation is greater for non-conservative amino acid substitutions; therefore, there is a higher probability of their having a more detrimental overall effect.

Nonsense mutations

The outcome is expected to be a truncated protein molecule. This may, or may not, get as far as entering the circulation, but even if it does, truncation is highly likely to have made it defunct. It is therefore not surprising that nonsense mutations are associated with severe haemophilia. Exon skipping is a further possibility arising from a nonsense mutation[86 87] and this is also extremely detrimental: an in frame skip will result in a protein lacking the amino acids encoded by the skipped exon, an out of frame skip will result in a frame shift.

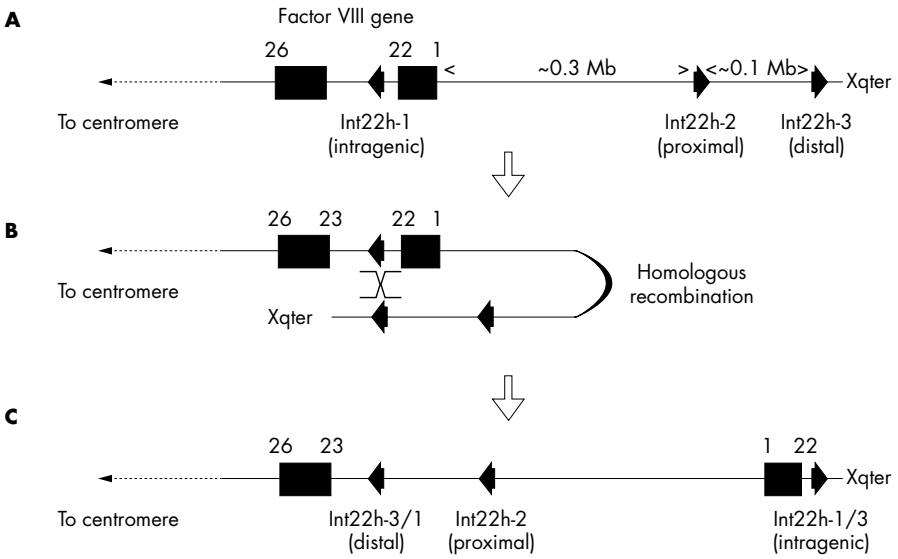


Figure 8 The intron 22 inversion of the factor VIII gene. (A) The normal configuration of the factor VIII gene and int22h homologues. (B) The distal or the proximal int22h homologue aligns with the intragenic int22h homologue (distal shown here). Homologous recombination takes place. (C) Linearisation of the X chromosome gives a completely disrupted factor VIII gene: exons 1 to 22 are displaced towards the telomere and are orientated in a direction opposite to their normal orientation. A contiguous factor VIII gene transcript running from exon 1 to 26 is no longer possible and severe haemophilia A results.

mRNA splice site mutations

Mutations that destroy or create mRNA splice sites are associated with variable severity of haemophilia; this very much depends on whether some correct transcripts can be processed (mild to moderate disease) or whether there is a complete loss of correct mRNA processing (severe disease). Exon skipping is a further possible consequence of splice site mutations, the outcome of which depends on whether the skip is in frame or results in a frame shift.[88]

Point mutations can be grouped according to whether or not they are classic CG site transitions. CG sites are hypermutable and give rise to C to T transitions or G to A transitions.[89] The mechanism underlying these transitions is understood[89] and is a major mechanism of human gene mutation. This is clearly exemplified in the haemophilias, in which approximately 30% of all distinct point mutations arise at CG sites. The remaining 70% of distinct point mutations do not occur at CG sites and may arise—for example, as a result of nucleotide misincorporation during DNA replication.

Point mutations may also be grouped according to whether or not they are recurrent. Recurrent mutations occur in both haemophilia A and B and are principally found at CG sites. The hypermutability of such sites underlies the independent recurrence of their transition events. There are rare examples of recurrent mutations that do not occur at CG sites. These are generally explained by a founder event that has multiplied in successive generations. Evidence for this may include confined geographical distribution and identical polymorphic haplotype in all affected genes, irrespective of geographical distribution.[90 91] Founder mutations typically cause mild/moderate disease; a mutation causing severe disease has a low probability of survival over many generations.

Deletions

Deletions of the factor VIII and factor IX gene include whole gene deletions, partial gene deletions at the 5′ end, 3′ end, or within the gene, and microdeletions of one to several base pairs (gross deletions are compiled in several reviews[14 28 92 93]). They do not appear to cluster to a specific region of either gene, but appear to be randomly distributed. A deletion has a high probability of destroying genetic function, removing a chunk of protein, or introducing a frame shift, all of which are extremely detrimental. Not surprisingly therefore, deletions are generally associated with severe disease.

Insertions

Insertions have been characterised in both the factor VIII and factor IX genes. Like deletions, these can be gross or as small as one or several nucleotides, and gene function or the gene product can be very adversely affected (gross insertions are compiled in several reviews[14 28 92 93]).

> "Deletions are generally associated with severe disease"

Rearrangements/inversions

These are rare gene defects in the haemophilias, except for an inversion involving intron 22 of the factor VIII gene, which causes severe haemophilia A, and which is found in 40–50% of patients with severe disease.[94 95]

The intron 22 inversion arises through homologous recombination between int22h-1 and int22h-2 (proximal) or int22h-3 (distal) during meiosis[94 95] (fig 8A–C). After homologous alignment, recombination results in the 5′ portion of the factor VIII gene being juxtaposed to the telomeric DNA, whereas the relevant extragenic homologue becomes juxtaposed to the 3′ portion of the factor VIII gene (fig 8B). In the re-linearised chromosome, the result is an inversion of the sequence between int22h-1 and int22h-2 or int22h-3, depending on which extragenic copy aligned (fig 8C). The factor VIII gene is completely disrupted: introns 1 to 22 are moved away from their normal context and their orientation is inverted. Severe haemophilia A results.

The intron 22 inversion is principally an error of DNA replication during spermatogenesis and not oogenesis[96]: in males, the absence of a second X chromosome for homologous alignment during meiosis would favour intrachromosomal alignments where they are possible. Thus, in haemophilia A families in which the intron 22 inversion is the causative gene defect, the defect can often be shown to have originated in an unaffected male relative, and in sporadic cases, this is often the patient's grandfather on his mother's side of the family.[96] The distal inversion is more common than the proximal inversion,[96] which may be explained by the greater genetic distance between the factor VIII gene and the distal int22h-3

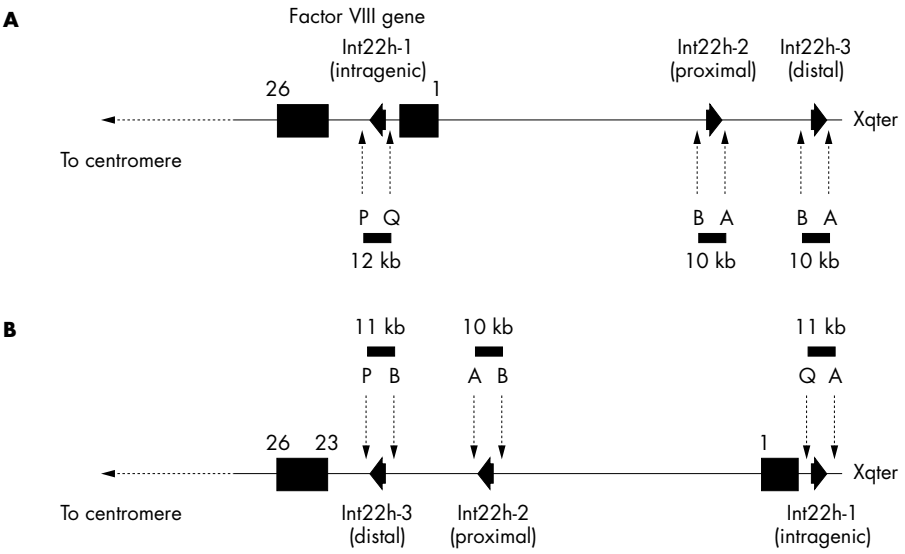


Figure 9 Detection of the intron 22 inversion of the factor VIII gene using long distance PCR (LD-PCR).[97] Primers P and Q hybridise specifically to sequences within the factor VIII gene and flank int22h-1. Primers A and B hybridise specifically to sequences flanking int22h-2 and int22h-3. The normal factor VIII gene therefore gives the LD-PCR products P + Q (12 kb) and A + B (10 kb). The intron 22 inversion results in the hybridisation site for primer Q being displaced towards the telomere, whereas the hybridisation site for primer B from the distal or proximal int22h homologue (depending on which was involved in the homologous recombination leading to the inversion, distal shown here) is brought into context with primer P. The inverted gene therefore gives the LD-PCR products P + B (11 kb), A + B (10 kb), and Q + A (11 kb). An A + B product is always obtained because one copy of the int22h homologues is unaffected by the inversion.

homologue facilitating the formation of the loop required for alignment to take place.

The intron 22 inversion was, until recently, diagnosed using the labour intensive approach of Southern blot analysis (the DNA fragmentation pattern obtained for the normal gene differs from that obtained for an inverted gene, and the patterns for the distal and proximal inversions also differ).[94 95] This has now been superseded by an approach using LD-PCR (fig 9).[97] This impressive strategy detects normal and inverted genes; however, it does not distinguish proximal inversion from distal inversion.

Although most intron 22 inversions involve int22h-1 and 2/3, atypical inversions have been reported.[94 98 99] Detailed analysis of one of these revealed a third, truncated copy of int22h between the factor VIII gene and int22h-2, and this was involved in the inversion.[98] These atypical inversions give anomalous patterns using Southern blot. They may fail to be detected using LD-PCR and this has implications for diagnosis using this approach. Fortunately, they appear to be rare.

MOLECULAR PATHOLOGY: CAUSE–EFFECT RELATIONS

The following detailed examples offer a glimpse into the many and varied mechanisms through which mutations bring about their effects. Mutations can exert their effects at any stage from genetic function right through to protein function, and often they reveal important properties of these functions that might otherwise have escaped realisation. The investigation of cause–effect relations is therefore of considerable value with potentially high rewards.

Nonsense mutations at CGA (arginine) codons

CG transitions can give rise to missense or nonsense mutations, depending upon the codon in which the CG site resides and whether C to T or G to A transition has occurred. C to T transition is particularly damaging at arginine codons of the type CGA: the result is the translation stop codon TGA (a nonsense mutation). As discussed above, nonsense mutations are particularly detrimental and are associated with severe disease. The factor VIII open reading frame has 70 CG sites, 12 of which reside in arginine codons of the type CGA (fig 10A). The factor IX open reading frame has 20 CG sites, six of which reside in CGA codons (fig 10B). C to T transition resulting in nonsense mutations has been reported at all of these codons and is associated with severe haemophilia, either haemophilia A[100–104] or B[105–107]. As would be expected from the hypermutability of CG sites, many of these nonsense mutations are recurrent among patients with haemophilia.

Missense mutations at activation cleavage sites

The proteolytic cleavages that activate factor VIII and factor IX occur after arginine residues located at positions 372 and 1689 in factor VIII (fig 3C), and 145 and 180 in factor IX (fig 4C). The codons for these arginine residues all contain the dinucleotide CG; transition of C to T or G to A within any of these codons causes haemophilia, either haemophilia A[103 104 108] or B.[107 109 110] Either transition results in a missense mutation that causes the relevant activation cleavage site arginine to be replaced by another amino acid (fig 11A, B). This compromises the activation cleavage at that site. Normal activation can therefore take place at only one of the two necessary sites and the result is a factor VIII or factor IX molecule with decreased function. Partial activity may remain, depending upon the activation cleavage site that is lost and the amino acid substitution. For example, factor IX cleaved at the β cleavage site alone retains partial activity, whereas factor IX cleaved at the α cleavage site alone is inactive; β cleavage is essential for unmasking the serine protease function of the catalytic domain. Thus, mutant proteins such as factor IX Chapel Hill[111] (histidine 145) retain some residual activity upon cleavage at the β bond, despite loss of cleavage at the α bond. The same is true for the cysteine 145 mutation (factor IX Albuquerque[112]). Mild haemophilia B results. In contrast, factor IX Hilo[113] (glutamine 180) cannot be activated at the β cleavage site, which is essential for the function of the enzyme, and severe haemophilia B results.

Mutations affecting factor VIII binding to von Willebrand factor

In plasma, factor VIII circulates bound to von Willebrand factor and this interaction stabilises factor VIII.[16 17] The regions of factor VIII that were found to be important for this interaction

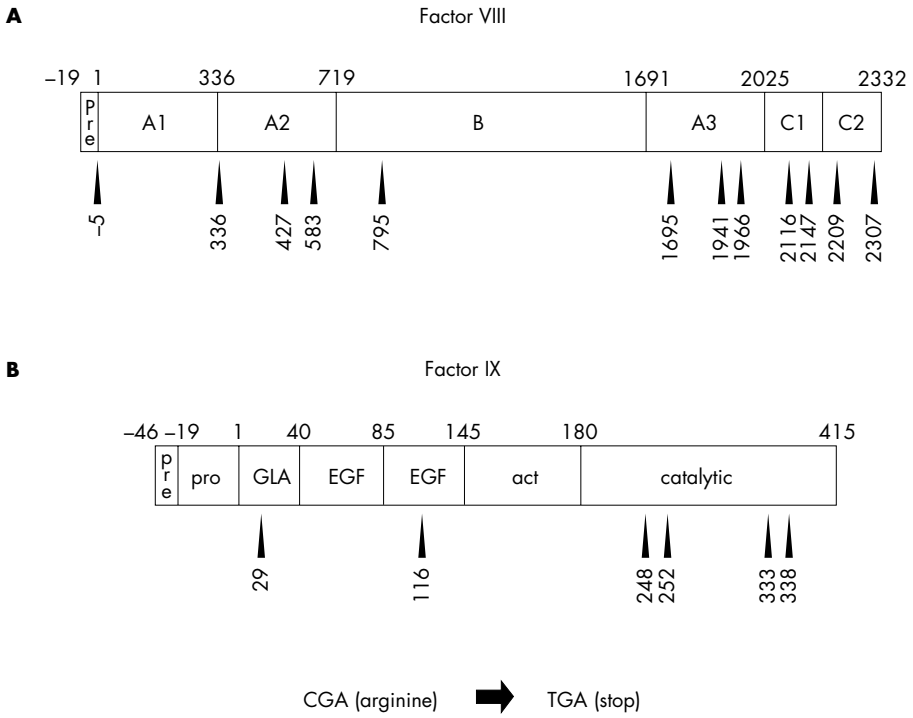


Figure 10 Arginine residues encoded by CGA within (A) the factor VIII and (B) the factor IX protein sequences. Nonsense mutations are brought about by C to T transition within the CG site of these codons. These are recurrent gene defects within haemophilia populations.

are the N-terminal part of the A3 domain, the C2 domain and, more recently, the C1 domain.[114–117] Mutations that interfere with factor VIII binding to von Willebrand factor illustrate how different mechanisms can affect the same property.

Within the A3 domain, sulphation of tyrosine 1680 is crucial for factor VIII interaction with von Willebrand factor.[118] Missense mutations resulting in the replacement of this tyrosine residue result in decreased factor VIII binding to von Willebrand factor[104 118 119] Similarly, retention of the tyrosine residue, but prevention of sulphation, also decreases factor VIII binding to von Willebrand factor.[120] Clearly, the loss of sulphation is crucial, whether or not this is accompanied by structural/folding changes brought about by a mutant amino acid substitution.

Several mutations have been identified within the C1 domain that decrease factor VIII binding to von Willebrand factor.[117 121] Some of these occur in residues that are conserved in the discoidin family of proteins; therefore, they are more likely to play a structural role in the C1 domain than to be involved directly in intermolecular interactions with von Willebrand factor. The structural perturbation that ensues compromises the interaction between the two proteins. Examples of such mutations are serine 2119 to tyrosine, asparagine 2129 to serine, arginine 2150 to histidine, and proline 2153 to glutamine.[117] However, none of these mutations sheds light on the mechanism by which C1 influences the affinity of factor VIII for von Willebrand factor: residues within the C1 domain may directly interact with von Willebrand factor residues;

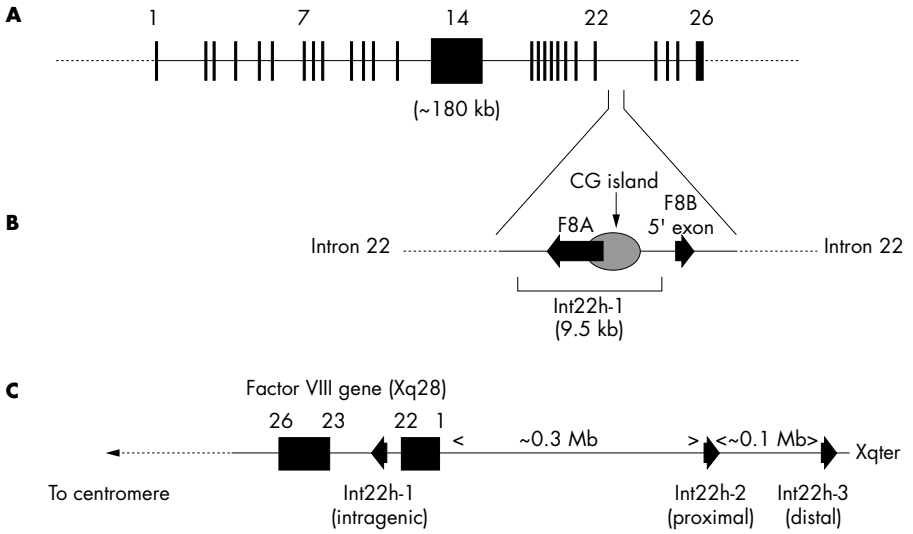


Figure 11 Mutations as a result of CG transitions at the codons encoding the activation cleavage site arginine residues of (A) factor VIII and (B) factor IX.

alternatively, mutations in the C1 domain may have an indirect effect through perturbation of folding or structure.

Besides mutations within the factor VIII gene that affect interaction with von Willebrand factor, there are also mutations of the von Willebrand factor gene that decrease the affinity of von Willebrand factor for factor VIII (von Willebrand disease type 2N).[122] Von Willebrand disease type 2N is typified by decreased factor VIII values, but shows an autosomal pattern of inheritance (the von Willebrand factor gene is located on chromosome 12 at 12p12–12pter). For obvious reasons, von Willebrand disease type 2N has also been referred to as pseudohaemophilia A.[123] There is an increasing list of mutations underlying von Willebrand disease type 2N. These are listed on the World Wide Web at URL http://mmg2.im.med.umich.edu/vWF/ (von Willebrand Factor (VWF) Database).

Mutations affecting factor VIII secretion

Factor VIII values may be decreased in the circulation by mutations that compromise its secretion. The haemophilia A missense mutations arginine 593 to cysteine and asparagine 618 to serine cause intracellular accumulation of factor VIII when expressed in cell culture.[124] This was concurrent with decreased factor VIII activity and antigen in the conditioned medium of transfected cells. Intracellular accumulation and decreased secretion of factor VIII possessing either of these mutations may explain the reduction of both factor VIII activity and antigen in plasma of patients with mild haemophilia A and these gene defects.

Mutations affecting factor IX binding to factor VIII

Initial evidence indicating that the protease domain of activated factor IX is involved in binding to activated factor VIII was provided by the inhibition of the interaction by a monoclonal antibody directed against the protease domain.[125] Subsequent antibody and mutation studies have revealed a complex story.

The factor IX protease domain contains a Ca^{2+} binding site in which glutamic acid residue 245 participates in binding the metal ion.[126] Substitution of glutamic acid 245 by valine (a haemophilia B mutation) results in a factor IX molecule that is rapidly cleaved between arginine 318 and serine 319 by activated factor XI.[127] This variant factor IX has a 15 times lower affinity for factor VIII before cleavage at arginine 318 and a 120 times lower affinity after cleavage.[127] Currently, it is believed that binding of Ca^{2+} to the protease domain protects factor IX from proteolysis after arginine 318, and this stabilises a neighbouring surface exposed helix (residues 330–338), which binds factor VIII.[84 127]

Mutations of eight of the nine residues in this helix (leucine 330 to arginine 338) are reported in the haemophilia B population and they probably cause haemophilia, at least in part, by directly interfering with factor IX–factor VIII interaction. In contrast, the haemophilia B mutation glutamic acid 245 to valine abolishes Ca^{2+} binding and this has an indirect effect on factor VIII binding, through destabilisation and possible structural effects, as discussed above.[128]

Mutations causing factor IX deficiency during anticoagulant treatment

Bleeding complications arising from oral anticoagulant treatment are not uncommon. An unusually selective decrease in factor IX procoagulant activity (1–3% of normal) has been observed in some patients receiving coumarin.[129] Normal factor IX activity was recovered on withdrawal of the anticoagulant. Two different missense mutations at alanine −10 were found to be implicated: alanine to valine and alanine to threonine. Alanine −10 lies within the pro-peptide domain, at a position that is essential for binding of the carboxylase that modifies glutamic acid residues in the adjacent GLA domain (fig 4C).[130 131] This carboxylase requires vitamin K as cofactor.[132 133] Coumarin derivatives interrupt the vitamin K cycle and lead to rapid depletion of the cofactor form of the vitamin.[134] Carboxylase activity is consequently decreased and this, in conjunction with a decreased affinity of the carboxylase for the mutant pro-domain,[131 135] results in the loss of GLA carboxylation and hence of factor IX activity. Presumably, in the absence of coumarin, normal vitamin K concentrations supply the carboxylase adequately and this permits appropriate processing of the GLA domain, even though the affinity of the carboxylase for the pro-domain may be decreased by the alanine −10 mutations.

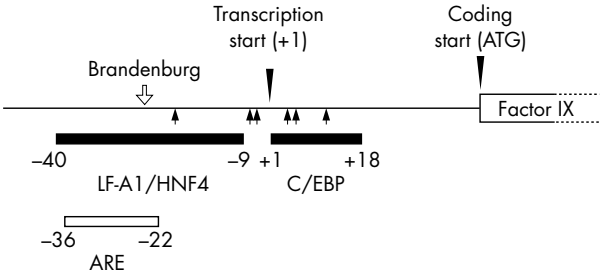


Figure 12 The promoter region of the human factor IX gene showing the androgen response element (ARE) and the regions implicated in binding LF-A1/HNF4, C/EBP. Coordinates are relative to the transcription start site (+1). Haemophilia B Leyden mutations (closed arrows) lie outside the ARE but within the LF-A1/HNF4 and C/EBP regions implicated in constitutive expression. In contrast, the promoter mutation G-26C (haemophilia B Brandenburg, open arrow) causes lifelong haemophilia B and lies within both the ARE and the −9 to −40 region associated with constitutive expression.

Mutations in the factor IX promoter and haemophilia B Leyden

The mutations discussed so far have affected either the factor VIII or factor IX coding sequence, with a consequent impact upon the translated protein. However, certain mutations within the promoter of the factor IX gene do not result in a qualitative change in the factor IX molecule. These mutations lie within sequences to which proteins involved in transcription bind. They disrupt or interfere with protein–DNA interactions in the promoter and so compromise the efficiency of transcription.

Some promoter mutations cause haemophilia B that is present throughout life.[136] However, others cause a variant of haemophilia B known as haemophilia B Leyden, which is characterised by severe factor IX deficiency at birth and through childhood, an increase in factor IX values from puberty through adolescence, and near normal factor IX values throughout adulthood.[137]

The molecular pathology of haemophilia B Leyden has been investigated intensively. The promoter of the factor IX gene contains a binding site for the CCAAT/enhancer binding protein[138] (C/EBP) and for the rat liver protein LF-A1/HNF4 (fig 12).[136] These proteins are involved in transcription; therefore, the binding sites may act as constitutive promoter elements for gene expression. The promoter also contains a candidate androgen response element (ARE), which overlaps the LF-A1/HNF4 element (fig 12).[136 139] This may act as a hormonally regulated promoter element for gene expression.

Before puberty, factor IX gene expression is regulated by constitutive promoter elements (C/EBP, LF-A1/HNF4, and possibly others). Mutations within these elements disrupt binding of the relevant proteins for transcription and haemophilia B results. At puberty, and thereafter, it is assumed that testosterone dependent transcription, mediated by the ARE, comes into play. Thus, provided that a promoter mutation does not lie within the ARE, amelioration of the haemophilia B is possible after puberty.

"Certain mutations within the promoter of the factor IX gene disrupt or interfere with protein–DNA interactions in the promoter and so compromise the efficiency of transcription"

In contrast, promoter mutations that cause life long haemophilia B, such as haemophilia B Brandenburg[136] (G-26C), have been found to lie within the region of overlap between the ARE and the LF-A1/HNF4 element and appear to interfere with both constitutive and hormonally regulated factor IX gene expression (fig 12).

Haemophilia A arising from mutations within the promoter of the factor VIII gene has not yet been demonstrated.

MUTATION DETECTION STRATEGIES

Studies aimed at defining the causative mutation within haemophilia cohorts have repeatedly shown that most patients possess a defect within the essential regions of the factor VIII or factor IX gene. These regions comprise the 5′ promoter sequence, the exons, the intron–exon boundaries, and the 3′ region around the polyadenylation signal sequence.

Therefore, the most direct strategy for mutation detection would be to amplify these regions from genomic DNA using PCR and screen the PCR products for a mutation. In the case of the factor IX gene, this is a straightforward task: the gene has only eight exons, the largest of which is less than 2 kb, so that fewer than 10 amplifications are required to cover the essential regions of the gene.[105]

In contrast, the large size and complexity of the factor VIII gene would necessitate almost 30 amplifications of genomic DNA to cover the essential regions. This is manageable and is the strategy used in some laboratories.[75 140–142] Alternatively, factor VIII mRNA can be used to screen exons 1 to 13 and 15 to 26.[143] The large exon 14 and the 5′ promoter and 3′ polyadenylation signal region can be screened from genomic DNA.[143] Factor VIII mRNA is not obtained from the accepted primary site of transcription (the liver), but from circulating white blood cells in which the transcript is believed to be present as a result of ectopic (leaky) expression.[143]

Various mutation detection techniques can be used to screen the PCR products from the factor VIII or factor IX genes.[75] One of the most informative is chemical mismatch cleavage,[144] the approach used for the characterisation of haemophilia B mutations within the Swedish population[145] and within the UK population.[146] It has also been used extensively in the current effort to characterise the mutations within the UK haemophilia A population.[147] Solid phase, fluorescence based chemical mismatch cleavage has provided a rapid route to screening the essential regions of the complex factor VIII gene in large numbers of patients.[148]

Direct nucleotide sequence analysis using automatic DNA sequencers is becoming more mainstream[142] and confident results can be expected for male DNA (hemizygous). Sequencing should be interpreted cautiously for female DNA because heterozygosity may fail to show in the sequencing data. Multiplex amplification of all of the essential regions of the factor IX gene in a single PCR, followed by sequencing,[149] represents a further step forward and one that could be applied to the factor VIII gene.

MUTATIONS AND THE RISK OF INHIBITORS

The production of neutralising antibodies in response to infused factor VIII or factor IX has always been of considerable interest, principally because it is a major complication of replacement treatment.[150 151] It is also potentially a major complication of gene therapy; therefore, insight into the factors that predispose towards inhibitor production would be of considerable value. Evidently, the mutation underlying the haemophilia is important[141 150–153]: mutations associated with the absence of a gene product, such as deletions or nonsense mutations, confer a high risk for inhibitor production; mutations associated with the presence of a gene product (even very low amounts of the protein) confer a low risk for inhibitor production. However, in reality, the situation is not so clear cut. Among patients with identical mutations, some may produce inhibitors and others may not.[141 154] Clearly other factors are implicated. The possibility that these may include the genotype at the (MHC) locus has been investigated in patients with severe haemophilia A, with and without the intron 22 inversion.[155 156] Only a weak association between human MHC (HLA) class II genotype and the development of inhibitor antibodies against factor VIII was obtained; this was slightly more pronounced in patients with the inversion than in those without.

Subtle effects may attenuate the relation between a haemophilia mutation and the risk of inhibitor production. For example, one would expect small insertions or deletions resulting in an alteration of the protein translation reading frame (frameshift mutations) to confer a high risk of inhibitor production. However, several groups have found the absence of inhibitors among patients with haemophilia A possessing a small deletion or insertion within a run of the same nucleotide.[141 157] A proposed explanation is that transcriptional or translational errors may lead, fortuitously, to the occasional correction of the reading frame and this may lead to sufficient endogenous factor VIII to explain immune tolerance.[141 157 158] Patients with a small deletion or insertion not within a run of the same nucleotide may not be so lucky.[141]

It therefore appears, at present, a difficult task, even with the knowledge of a patient's causative mutation, to predict an inhibitor response. Fortunately, however, inhibitors are a rare phenomenon among patients with haemophilia (apparently rarer in haemophilia B than haemophilia A). As one might expect, they are principally found within the severe group (high probability of the absence of a gene product), where they are found in 10–15% of individuals.[150 153]

FACTOR VIII AND FACTOR IX TURNOVER

The steady state concentrations of factor VIII and factor IX in the circulation reflect a balance of synthesis and removal. The clearance mechanism for both factors appears to be mediated, at least in part, by the low density lipoprotein receptor related protein (LRP), a liver multiligand endocytic receptor. Although the interaction of factor VIII with LRP is not disputed, there is some controversy concerning the region of factor VIII involved in binding and the effect of von Willebrand factor on the interaction. Factor VIII may bind LRP via a region within the light chain[159] and this may be inhibited by von Willebrand factor,[159] which also binds factor VIII via the light chain.[118] This could then explain the beneficial effect of von Willebrand factor on the in vivo survival of factor VIII.[16 17] Alternatively, factor VIII may bind LRP via the A2 domain in an interaction that is not influenced by von Willebrand factor.[160] In vivo studies in von Willebrand factor deficient mice suggest that the accelerated clearance of factor VIII seen in the absence of von Willebrand factor may be the result of LRP mediated factor VIII catabolism.[161] Clearly, LRP mediated removal of factor VIII appears to be real, and whether or not this is modulated by von Willebrand factor remains to be clarified.

LRP may also be involved in the catabolism of factor IX. Factor IX interaction with LRP appears to require activation of the coagulation factor by activated factor XI.[162] Delivery of activated factor IX to LRP appears to require proteoglycans on the cell surface.[162] In vitro, degradation of factor IX by cells deficient in either LRP or proteoglycans was significantly decreased compared with cells expressing these proteins. The relevance of these findings has yet to be established in vivo, although they suggest that one route to catabolism of activated factor IX may be via a proteoglycan on the cell surface, which delivers factor IX to LRP, which then targets the factor IX to the intracellular degradation pathway.

"Insights gained from mutations and their effects—the molecular pathology of disease—have revealed both subtle and major complexities at the genetic and protein levels"

CONCLUDING COMMENTS

The exploration of the factor VIII and factor IX genes and their gene products over the past 15 years has provided enormous amounts of information and has expanded into many and varied territories. To review all of these comprehensively would require several volumes; therefore, this review has necessarily been restricted. The areas covered reveal the ingenuity of the research community (scientist and clinician alike) and humankind's incessant struggle for knowledge and understanding. As each question is answered, and new questions arise, the resolution of our understanding sharpens.

Insights gained from mutations and their effects—the molecular pathology of disease—have revealed both subtle and major complexities at the genetic and protein levels. Similarly, insights gained from polymorphisms and their ethnic variations hint at a complex evolution awaiting further exploration. Finally, insights gained from the investigation of patients' genetics and the aetiology of the inhibitor complication are paving the path for gene therapy in previously untreated patients. The molecular biology of the haemophilias has provided food for research and food for thought, and the feast seems certain to continue into the distant future.

APPENDIX A

DEFINITIONS OF POLYMORPHISMS FOUND IN THE FACTOR VIII AND FACTOR IX GENES, AND PROPERTIES RELEVANT TO LINKAGE STUDIES IN THE HAEMOPHILIAS

Single nucleotide polymorphisms

Single nucleotide polymorphisms (SNPs) are precisely what their name suggests: nucleotide positions at which different nucleotides are present in the gene in the normal population. Typically, they are biallelic (one of two nucleotides is present within any given gene).

Many SNPs fortuitously reside within recognition sequences for restriction endonucleases and this provides a convenient means by which to genotype them. Such SNPs are better known as restriction fragment length polymorphisms (RFLPs).

SNPs can be intronic or exonic. Those within exons can either be silent in that they do not change the amino acid encoded by the codon in which they occur, or they can change a codon so that a different amino acid is specified. The latter give rise to amino acid polymorphisms within the protein sequence.

Microsatellites

Microsatellites are sequences built up from multiple repeats of a given sequence motif and are polymorphic in that the number of repeats present within the gene varies in the normal population. They are typically multiallelic. The human genome contains a diversity of microsatellites, just one example of which is the sequence (CA)n, where "n" represents a variable number of repeats among different genes in the normal population.

Frequency

The alleles of SNPs and microsatellites occur with specific frequencies in the general population. In the case of biallelic SNPs, the equal occurrence of both alleles would result in a relative frequency of 0.5 for each. However, in reality, such equivalence in frequency between both alleles is rarely seen and a preponderance of one allele over the other is more usual—for example, (0.1/0.9) or (0.4/0.6) (allele 1/allele 2).

In the case of microsatellites, several alleles are usually present within the general population and their relative frequencies are not equal. One, two, or three alleles may predominate and other alleles may be more rare. The existence of several alleles is an important property of microsatellites in the context of linkage analysis (see later).

The allele frequencies of any polymorphism may vary between people from different racial backgrounds and, indeed, some loci may be polymorphic within one racial group but non-polymorphic within another.

Heterozygosity, homozygosity, and hemizygosity

Females with a normal karyotype (46:XX) have two copies of each polymorphism located on the X chromosome and this, of course, includes those polymorphic loci found within the factor VIII and factor IX genes. For a given polymorphic locus within these genes, a female may have identical alleles (homozygous) or she may have one allelic variant in one gene and a different allelic variant in the other (heterozygous).

The probability of heterozygosity can be estimated from the allele frequencies in the general population. For example, a biallelic locus for which the alleles occur at frequencies of 0.1/0.9 can be expected to show the following distribution of genotypes:

Homozygous (allele 1/allele 1): $0.1 \times 0.1 = 0.01$.

Heterozygous (allele 1/allele 2): $0.1 \times 0.9 = 0.09$.

Heterozygous (allele 2/allele 1): $0.1 \times 0.9 = 0.09$.

Homozygous (allele 2/allele 2): $0.9 \times 0.9 = 0.81$.

Thus, the expected frequency of heterozygosity (the rate of heterozygosity) is $0.09 + 0.09 = 0.18$. Another way of interpreting this is that 18% of females may be expected to be heterozygous at that locus. An identical calculation done using a biallelic locus with allele frequencies of 0.5/0.5 yields a rate of heterozygosity of 0.5. This represents the maximum possible rate of heterozygosity for a biallelic locus and it can be interpreted to mean that 50% of females can be expected to be heterozygous at that locus.

Microsatellites yield much higher rates of heterozygosity because there are greater numbers of allelic variants and co-inheritance of identical alleles is intrinsically less probable than for a biallelic locus. For example, in a microsatellite with three allelic variants showing frequencies of 0.1 (allele 1), 0.6 (allele 2), and 0.3 (allele 3), the expected distribution of genotypes is as follows:

Homozygous (allele 1/allele 1): $0.1 \times 0.1 = 0.01$.

Homozygous (allele 2/allele 2): $0.6 \times 0.6 = 0.36$.

Homozygous (allele 3/allele 3): $0.3 \times 0.3 = 0.09$.

Heterozygous (allele 1/allele 2): $0.1 \times 0.6 = 0.06$.

Heterozygous (allele 2/allele 1): $0.6 \times 0.1 = 0.06$.

Heterozygous (allele 1/allele 3): $0.1 \times 0.3 = 0.03$.

Heterozygous (allele 3/allele 1): $0.3 \times 0.1 = 0.03$.

Heterozygous (allele 2/allele 3): $0.6 \times 0.3 = 0.18$.

Heterozygous (allele 3/allele 2): $0.3 \times 0.6 = 0.18$.

From these calculations, the expected rate of heterozygosity is $0.06 + 0.06 + 0.03 + 0.03 + 0.18 + 0.18 = 0.54$. This exceeds the 0.5 maximum for biallelic loci and yet it has only involved a locus with three alleles, one of which is of very low frequency (allele 1, 0.1). In practice, microsatellites have many more than three alleles and it is not unusual to observe rates of heterozygosity approaching 0.8.

A male with a normal karyotype (46:XY) only has one factor VIII gene and factor IX gene, therefore he will have only one copy of a given polymorphic locus and only one allele will be present (hemizygous).

Informativity and non-informativity

The outcome of a linkage study cannot be predicted at the outset: the data must be generated and then interpreted to reveal whether the polymorphic loci investigated have permitted a defective (or normal) gene to be tracked through the family. Any polymorphic loci that permit gene tracking from

one individual to another within a family are said to be informative for that individual. Conversely, those loci that do not permit the gene to be tracked from one individual to another are non-informative for that individual.

Evidently, heterozygosity is associated with informativity: the two genes can be discerned by virtue of their different polymorphic alleles and this permits them to be tracked separately. Homozygosity is associated with non-informativity because the two genes cannot be distinguished. However, homozygosity can be helpful in establishing haplotypes from genotypes (see below).

Genotypes and haplotypes

Analysis of polymorphic loci to determine the alleles present generates a genotype for each individual in a family. This must then be resolved into haplotypes (the alleles in one gene separated from those in the other). In the case of the haemophilias, this is made easier by virtue of hemizygosity in males: the genotype is the haplotype in the case of hemizygosity. This may then permit haplotypes in females to be established by subtraction of a known haplotype from the genotype to give the second haplotype.

Homozygosity in females also permits the determination of haplotypes: if a female is homozygous, both genes must have identical haplotypes for the polymorphic loci examined, and this may permit further haplotypes to be resolved within the family.

In phase

The haplotype found in the factor VIII or factor IX gene of the male haemophiliac within a family is said to be in phase with haemophilia: it identifies the defective factor VIII or factor IX gene. It is worth stressing that this haplotype undoubtedly exists in the normal population where it is not associated with a defective gene, it just so happens that, in the family being investigated, the defective gene has this haplotype. Thus, discovery of the same haplotype in other members of the family does not necessarily define the presence of the defective gene. Inheritance of the defective gene has to be shown directly by analysis of relevant family members.

Linkage disequilibrium

Polymorphisms within a given gene are necessarily linked. When a certain allele at one locus always occurs in conjunction with a certain allele at another locus, the loci are said to show complete linkage disequilibrium. When the alleles at one locus occur entirely independently of the alleles at a second locus, the loci are said to show complete linkage equilibrium. Between these two extremes there is a continuous scale of linkage disequilibrium.

If two loci show complete linkage disequilibrium, the information obtained at one will be identical to that obtained at the other and no additional information will be gained by examination of both. If one locus is homozygous and non-informative, the other one will be also, so that informativity is not increased by investigation of both loci. For this reason it may seem redundant to investigate loci showing linkage disequilibrium. However, examination of both does permit corroboration of the information obtained from either. This can be valuable, lending confidence to the data generated in the study. Loci that are in linkage disequilibrium may therefore not increase the informativity of a study very much, but they do provide a double check of the data, and this may be invaluable.

Implications of these properties of polymorphisms

The goal of a linkage study in the haemophilias is to establish carrier status and/or offer prenatal diagnosis with confidence. This requires an informative genetic result, and this should be derived preferably from more than one informative polymorphic locus. For this to be achieved, the best loci to examine are those with the highest rates of heterozygosity and the lowest degree of linkage disequilibrium.

APPENDIX B

DATABASE ENTRIES AT THE NATIONAL CENTRE FOR BIOTECHNOLOGY INFORMATION (URL HTTP://WWW.NCBI.NLM.NIH.GOV/)

Factor VIII (F8C)

- Gene: NT_025284
- mRNA: NM_000132
- Protein: NP_000123
- Locus Link: LocusID 2157

F8A

- Gene: NT_025307
- mRNA: NM_012151
- Protein: NP_036283
- Locus Link: LocusID 8263

F8B

- Gene: NT_025284
- mRNA: M90707
- Protein: AAA58466
- Locus Link: LocusID 2157

F8 Int22h-1

- DNA: X86011 and X86012

Factor IX (F9)

- Gene: K02402
- mRNA: NM_000133
- Protein: NP_000124
- Locus Link: LocusID 2158

REFERENCES

1 **Gitschier J**, Wood WI, Goralka TM, *et al.* Characterization of the human factor VIII gene. *Nature* 1984;**312**:326–30.
2 **Yoshitake S**, Schach BG, Foster DC, *et al.* Nucleotide sequence of the gene for human factor IX (antihemophilic factor B). *Biochemistry* 1985;**24**:3736–50.
3 **Drayna D**, White R. The genetic linkage map of the human X chromosome. *Science* 1985;**230**:753–8.
4 **Patterson M**, Schwartz C, Bell M, *et al.* Physical mapping studies on the human X chromosome in the region Xq27–Xqter. *Genomics* 1987;**1**:297–306.
5 **Wollina K**, Bowen DJ, Syrbe G, *et al.* Female twins with severe Christmas disease (hemophilia B). *Thromb Haemost* 1993;**70**:774–6.
6 **Ingerslev J**, Schwartz M, Lamm LU, *et al.* Female haemophilia A in a family with seeming extreme bidirectional lyonization tendency: abnormal premature X-chromosome inactivation? *Clin Genet* 1989;**35**:41–8.
7 **Orstavik KH**, Scheibel E, Ingerslev J, *et al.* Absence of correlation between X chromosome inactivation pattern and plasma concentration of factor VIII and factor IX in carriers of haemophilia A and B. *Thromb Haemost* 2000;**83**:433–7.
8 **Levinson B**, Kenwrick S, Lakich D, *et al.* A transcribed gene in an intron of the human factor VIII gene. *Genomics* 1990;**7**:1–11.
9 **Levinson B**, Kenwrick S, Gamel P *et al.* Evidence for a third transcript from the human factor VIII gene. *Genomics* 1992;**14**:585–9.
10 **Naylor JA**, Buck D, Green P, *et al.* Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Hum Mol Genet* 1995;**4**:1217–24.
11 **Levinson B**, Bermingham JR, Jr, Metzenberg A, *et al.* Sequence of the human factor VIII-associated gene is conserved in mouse. *Genomics* 1992;**13**:862–5.
12 **Wood WI**, Capon DJ, Simonsen CC, *et al.* Expression of active human factor VIII from recombinant DNA clones. *Nature* 1984;**312**:330–7.
13 **Vehar GA**, Keyt B, Eaton D, *et al.* Structure of human factor VIII. *Nature* 1984;**312**:337–42.
14 **Tuddenham EGD**, Cooper DN. Factor VIII and haemophilia A. *Oxford Monographs on Medical Genetics* 1994;**25**:19–76.
15 **Eaton D**, Rodriguez H, Vehar GA. Proteolytic processing of human factor VIII. Correlation of specific cleavages by thrombin, factor Xa, and

activated protein C with activation and inactivation of factor VIII coagulant activity. *Biochemistry* 1986;**25**:505–12.
16 **Weiss HJ**, Sussman II, Hoyer LW. Stabilization of factor VIII in plasma by the von Willebrand factor. Studies on posttransfusion and dissociated factor VIII and in patients with von Willebrand's disease. *J Clin Invest* 1977;**60**:390–404.
17 **Tuddenham EG**, Lane RS, Rotblat F, *et al*. Response to infusions of polyelectrolyte fractionated human factor VIII concentrate in human haemophilia A and von Willebrand's disease. *Br J Haematol* 1982;**52**:259–67.
18 **Hamer RJ**, Kodam JA, Beeser-Viser NH, *et al*. The effect of thrombin on the complex between factor VIII and von Willebrand factor. *Eur J Biochem* 1987;**167**:253–9.
19 **Saenko EL**, Shima M, Sarafanov AG. Role of activation of the coagulation factor VIII in interaction with vWf, phospholipid, and functioning within the factor Xase complex. *Trends Cardiovasc Med* 1999;**9**:185–92.
20 **Anson DS**, Choo KH, Rees DJ, *et al*. The gene structure of human anti-haemophilic factor IX. *EMBO J* 1984;**3**:1053–60.
21 **Di Scipio RG**, Kurachi K, Davie EW. Activation of human factor IX (Christmas factor). *J Clin Invest* 1978;**61**:1528–38.
22 **Walsh PN**, Bradford H, Sinha D, *et al*. Kinetics of the factor Xla catalyzed activation of human blood coagulation factor IX. *J Clin Invest* 1984;**73**:1392–9.
23 **Jesty J**, Morrison SA. The activation of factor IX by tissue factor–factor VII in a bovine plasma system lacking factor X. *Thromb Res* 1983;**32**:171–81.
24 **Osterud B**, Rapaport SI. Activation of factor IX by the reaction product of tissue factor and factor VII: additional pathway for initiating blood coagulation. *Proc Natl Acad Sci U S A* 1977;**74**:5260–4.
25 **Foster PA**, Zimmerman TS. Factor VIII structure and function. *Blood Rev* 1989;**3**:180–91.
26 **Giannelli F**. Factor IX. *Baillieres Clin Haematol* 1989;**2**:821–48.
27 **Lenting PJ**, van Mourik JA, Mertens K. The life cycle of coagulation factor VIII in view of its structure and function. *Blood* 1998;**92**:3983–96.
28 **Giannelli F**, Green PM. The molecular basis of haemophilia A and B. *Baillieres Clin Haematol* 1996; **9**:211–28.
29 **Jeffreys AJ**, Wilson V, Thein SL. Hypervariable "minisatellite" regions in human DNA. *Nature* 1985;**314**:67–73.
30 **Gill P**, Jeffreys AJ, Werrett DJ. Forensic application of DNA "fingerprints". *Nature* 1985;**318**:577–9.
31 **Paabo S**. Human evolution. *Trends Cell Biol* 1999;**9**:M13–16.
32 **Foley R**. The context of human genetic evolution. *Genome Res* 1998;**8**:339–47.
33 **Mibashan RS**, Peake IR, Nicolaides KH. Prenatal diagnosis of hemostatic disorders. *Methods Haematol* 1989;**21**:64–107.
34 **Peake IR**, Lillicrap DP, Boulyjenkov V, *et al*. Report of a joint WHO/WFH meeting on the control of haemophilia: carrier detection and prenatal diagnosis. *Blood Coagul Fibrinolysis* 1993;**4**:313–44.
35 **Venter JC**, Adams MD, Myers EW, *et al*. The sequence of the human genome. *Science* 2001;**291**:1304–51.
36 **Lander ES**, Linton LM, Birren B, *et al*. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–21.
37 **Kogan S**, Gitschier J. Mutations and a polymorphism in the factor VIII gene discovered by denaturing gradient gel electrophoresis. *Proc Natl Acad Sci U S A* 1990;**87**:2092–6.
38 **Lalloz MR**, McVey JH, Pattinson JK, *et al*. Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet* 1991;**338**:207–11.
39 **Gitschier J**, Drayna D, Tuddenham EG, *et al*. Genetic mapping and diagnosis of haemophilia A achieved through a BclI polymorphism in the factor VIII gene. *Nature* 1985;**314**:738–40.
40 **Ahrens P**, Kruse TA, Schwartz M, *et al*. A new HindIII restriction fragment length polymorphism in the hemophilia A locus. *Hum Genet* 1987;**76**:127–8.
41 **Wion KL**, Tuddenham EG, Lawn RM. A new polymorphism in the factor VIII gene for prenatal diagnosis of hemophilia A. *Nucleic Acids Res* 1986;**14**:4535–42.
42 **Chan V**, Tong TM, Chan TP, *et al*. Multiple XbaI polymorphisms for carrier detection and prenatal diagnosis of haemophilia A. *Br J Haematol* 1989;**73**:497–500.
43 **Bowen DJ**, De Brasi CD, Larripa IB, *et al*. A new polymorphism in the human factor VIII gene: implications for linkage analysis in haemophilia A and for the evolution of int22h sequences. *Br J Haematol* 2000;**111**:544–8.
44 **Lalloz MRA**, McVey JH, Michaelides K, *et al*. Haemophilia A diagnosis by analysis of a novel dinucleotide tandem repeat sequence within the factor VIII gene. *Br J Haematol* 1992;**80**:3a.
45 **Antonarakis SE**, Waber PG, Kittur SD, *et al*. Hemophilia A. Detection of molecular defects and of carriers by DNA analysis. *N Engl J Med* 1985;**313**:842–8.
46 **Youssoufian H**, Phillips DG, Kazazian HH, Jr, *et al*. MspI polymorphism in the 3′ flanking region of the human factor VIII gene. *Nucleic Acids Res* 1987;**15**:6312.
47 **Rai HK**, Winship PR. A –793 G to A transition in the factor IX gene promoter is polymorphic in the Caucasian population. *Br J Haematol* 1996;**92**:501–3.
48 **Hay CW**, Robertson KA, Yong SL, *et al*. Use of a BamHI polymorphism in the factor IX gene for the determination of hemophilia B carrier status. *Blood* 1986;**67**:1508–11.
49 **Winship PR**, Nichols CE, Chuansumrit A, *et al*. An MseI RFLP in the 5′ flanking region of the factor IX gene: its use for haemophilia B carrier detection in Caucasian and Thai populations. *Br J Haematol* 1993;**84**:101–5.
50 **Winship PR**, Anson DS, Rizza CR, *et al*. Carrier detection in haemophilia B using two further intragenic restriction fragment length polymorphisms. *Nucleic Acids Res* 1984;**12**:8861–72.
51 **Driscoll MC**, Dispenzieri A, Tobias E, *et al*. A second BamHI DNA polymorphism and haplotype association in the factor IX gene. *Blood* 1988;**72**:61–5.
52 **Giannelli F**, Anson DS, Choo KH, *et al*. Characterisation and use of an intragenic polymorphic marker for detection of carriers of haemophilia B (factor IX deficiency). *Lancet* 1984;**1**:239–41.
53 **Camerino G**, Oberle I, Drayna D, *et al*. A new MspI restriction fragment length polymorphism in the hemophilia B locus. *Hum Genet* 1985;**71**:79–81.
54 **Graham JB**, Kunkel GR, Tennyson GS, *et al*. The Malmo polymorphism of factor IX: establishing the genotypes by rapid analysis of DNA. *Blood* 1989;**73**:2104–7.
55 **Sommer SS**, Tillotson VL, Vielhaber EL, *et al*. "Cryptic" dinucleotide polymorphism in the 3′ region of the factor IX gene shows substantial variation among different populations. *Hum Genet* 1994;**93**:357–8.
56 **Winship PR**, Rees DJ, Alkan M. Detection of polymorphisms at cytosine phosphoguanadine dinucleotides and diagnosis of haemophilia B carriers. *Lancet* 1989;**1**:631–4.
57 **Rubinsztein DC**, Amos W, Leggo J, *et al*. Microsatellite evolution—evidence for directionality and variation in rate between species. *Nat Genet* 1995;**10**:337–43.
58 **Jeffreys AJ**, Royle NJ, Wilson V, *et al*. Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* 1988;**332**:278–81.
59 **Graham JB**, Kunkel GR, Fowlkes DM, *et al*. The utility of a HindIII polymorphism of factor VIII examined by rapid DNA analysis. *Br J Haematol* 1990;**76**:75–9.
60 **Suehiro K**, Tanimoto M, Hamaguchi M, *et al*. Carrier detection in Japanese hemophilia A by use of three intragenic and two extragenic factor VIII DNA probes: a study of 24 kindreds. *J Lab Clin Med* 1988;**112**:314–18.
61 **Chan V**, Chan TK, Liu VW, *et al*. Restriction fragment length polymorphisms associated with factor VIII:C gene in Chinese. *Hum Genet* 1988;**79**:128–31.
62 **Bowen DJ**. An effective route for the isolation of human genetic loci exemplified using the BglI restriction fragment length polymorphism of the factor VIII gene. *Thromb Haemost* 2000;**83**:512–13.
63 **Janco RL**, Phillips JA, III, Orlando P, *et al*. Carrier testing strategy in haemophilia A. *Lancet* 1986;**1**:148–9.
64 **Howarth A**, Bowen DJ. Linkage analysis in haemophilia A: simultaneous genotyping of two polymorphisms of the human factor VIII gene using induced heteroduplex formation. *Haemophilia* 1998;**4**:812–19.
65 **McGraw RA**, Davis LM, Noyes CM, *et al*. Evidence for a prevalent dimorphism in the activation peptide of human coagulation factor IX. *Proc Natl Acad Sci U S A* 1985;**82**:2847–51.
66 **Smith KJ**, Thompson AR, McMullen BA, *et al*. Carrier testing in hemophilia B with an immunoassay that distinguishes a prevalent factor IX dimorphism. *Blood* 1987;**70**:1006–13.
67 **Wallmark A**, Ljung R, Nilsson IM. Determination of factor IX allotypes for carrier identification in haemophilia B. *Br J Haematol* 1987;**67**:427–32.
68 **Sarkar G**, Paynton C, Sommer SS. Segments containing alternating purine and pyrimidine dinucleotides: patterns of polymorphism in humans and prevalence throughout phylogeny. *Nucleic Acids Res* 1991;**19**:631–6.
69 **Kojima T**, Tanimoto M, Kamiya T, *et al*. Possible absence of common polymorphisms in coagulation factor IX gene in Japanese subjects. *Blood* 1987;**69**:349–52.
70 **Toyozumi H**, Kojima T, Matsushita T, *et al*. Diagnosis of hemophilia B carriers using two novel dinucleotide polymorphisms and Hha I RFLP of the factor IX gene in Japanese subjects. *Thromb Haemost* 1995;**74**:1009–14.
71 **de la SC**, Wu Q, Baas MJ, *et al*. Common intragenic and extragenic polymorphisms of blood coagulation factors VIII and IX are different in Chinese and Caucasian populations. *Clin Genet* 1990;**38**:434–40.
72 **Graham JB**, Kunkel GR, Egilmez NK, *et al*. The varying frequencies of five DNA polymorphisms of X-linked coagulant factor IX in eight ethnic groups. *Am J Hum Genet* 1991;**49**:537–44.
73 **Figueiredo MS**, Bowen DJ, Silva Junior WA, *et al*. Factor IX gene haplotypes in Brazilian blacks and characterization of unusual DdeI alleles. *Br J Haematol* 1994;**87**:789–96.
74 **Zhang M**, Chen SH, Scott CR, *et al*. The factor IX BamHI polymorphism: T-to-G transversion at the nucleotide sequence –561. The BamHI/MSPI haplotypes in blacks and Caucasians. *Hum Genet* 1989;**82**:283–4.
75 **Goodeve AC**. Laboratory methods for the genetic diagnosis of bleeding disorders. *Clin Lab Haematol* 1998;**20**:3–19.
76 **El Maarri O**, Oldenburg J, Caglayan SH. Intron 22-specific long PCR for the Xba I polymorphism in the factor VIII gene. *Br J Haematol* 1999;**105**:1120–2.
77 **De Brasi CD**, Bowen DJ, Collins PW *et al*. Specific analysis of the intron 22 XbaI polymorphism of the human factor VIII gene using long-distance PCR. *Br J Haematol* 1999;**107**:566–8.
78 **Hacia JG**, Fan JB, Ryder O, *et al*. Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat Genet* 1999;**22**:164–7.
79 **Wilgenbus KK**, Lichter P. DNA chip technology ante portas. *J Mol Med* 1999;**77**:761–8.

80 **Diamandis EP**. Sequencing with microarray technology—a powerful new tool for molecular diagnostics. *Clin Chem* 2000;**46**:1523–5.
81 **Pemberton S**, Lindley P, Zaitsev V, *et al*. A molecular model for the triplicated A domains of human factor VIII based on the crystal structure of human ceruloplasmin. *Blood* 1997;**89**:2413–21.
82 **Brandstetter H**, Bauer M, Huber R, *et al*. X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and hemophilia B. *Proc Natl Acad Sci U S A* 1995;**92**:9796–800.
83 **Mertens K**, Celie PH, Kolkman JA, *et al*. Factor VIII–factor IX interactions: molecular sites involved in enzyme–cofactor complex assembly. *Thromb Haemost* 1999;**82**:209–17.
84 **Bajaj SP**. Region of factor IXa protease domain that interacts with factor VIIIa: analysis of select hemophilia B mutants. *Thromb Haemost* 1999;**82**:218–25.
85 **Wacey AI**, Tuddenham EG. Mutation databases on the Web. *J Med Genet* 1998;**35**:529–33.
86 **Dietz HC**, Valle D, Francomano CA, *et al*. The skipping of constitutive exons in vivo induced by nonsense mutations. *Science* 1993;**259**:680–3.
87 **Liu HX**, Cartegni L, Zhang MQ, *et al*. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat Genet* 2001;**27**:55–8.
88 **Tavassoli K**, Eigel A, Wilke K, *et al*. Molecular diagnostics of 15 hemophilia A patients: characterization of eight novel mutations in the factor VIII gene, two of which result in exon skipping. *Hum Mutat* 1998;**12**:301–3.
89 **Duncan BK**, Miller JH. Mutagenic deamination of cytosine residues in DNA. *Nature* 1980;**287**:560–1.
90 **Green PM**, Montandon AJ, Ljung R, *et al*. Haplotype analysis of identical factor IX mutants using PCR. *Thromb Haemost* 1992;**67**:66–9.
91 **Thompson AR**, Bajaj SP, Chen SH, *et al*. "Founder" effect in different families with haemophilia B mutation. *Lancet* 1990;**335**:418.
92 **Tuddenham EGD**, Cooper DN. Factor IX and haemophilia B. *Oxford Monographs on Medical Genetics* 1994;**25**:78–111.
93 **Antonarakis SE**. Molecular genetics of coagulation factor VIII gene and hemophilia A. *Thromb Haemost* 1995;**74**:322–8.
94 **Lakich D**, Kazazian HH, Jr, Antonarakis SE, *et al*. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nat Genet* 1993;**5**:236–41.
95 **Naylor J**, Brinke A, Hassock S, *et al*. Characteristic mRNA abnormality found in half the patients with severe haemophilia A is due to large DNA inversions. *Hum Mol Genet* 1993;**2**:1773–8.
96 **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**:2206–12.
97 **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**:1458–9.
98 **Naylor JA**, Nicholson P, Goodeve A, *et al*. A novel DNA inversion causing severe hemophilia A. *Blood* 1996;**87**:3255–61.
99 **Yamazaki E**, Mohri H, Inaba H, *et al*. Variant of intron 22 inversions in the factor VIII gene in severe hemophilia A. *Blood Coagul Fibrinolysis* 1997;**8**:445–9.
100 **Youssoufian H**, Kazazian HH, Jr, Phillips DG, *et al*. Recurrent mutations in haemophilia A give evidence for CpG mutation hotspots. *Nature* 1986;**324**:380–2.
101 **Youssoufian H**, Antonarakis SE, Bell W, *et al*. Nonsense and missense mutations in hemophilia A: estimate of the relative mutation rate at CG dinucleotides. *Am J Hum Genet* 1988;**42**:718–25.
102 **Gitschier J**, Kogan S, Levinson B, *et al*. Mutations of factor VIII cleavage sites in hemophilia A. *Blood* 1988;**72**:1022–8.
103 **Pattinson JK**, Millar DS, McVey JH, *et al*. The molecular genetic analysis of hemophilia A: a directed search strategy for the detection of point mutations in the human factor VIII gene. *Blood* 1990;**76**:2242–8.
104 **Higuchi M**, Kazazian HH, Jr, Kasch L, *et al*. Molecular characterization of severe hemophilia A suggests that about half the mutations are not within the coding regions and splice junctions of the factor VIII gene. *Proc Natl Acad Sci U S A* 1991;**88**:7405–9.
105 **Green PM**, Bentley DR, Mibashan RS, *et al*. Molecular pathology of haemophilia B. *EMBO J* 1989;**8**:1067–72.
106 **Koeberl DD**, Bottema CD, Sarkar G, *et al*. Recurrent nonsense mutations at arginine residues cause severe hemophilia B in unrelated hemophiliacs. *Hum Genet* 1990;**84**:387–90.
107 **Green PM**, Montandon AJ, Bentley DR, *et al*. The incidence and distribution of CpG→TpG transitions in the coagulation factor IX gene. A fresh look at CpG mutational hotspots. *Nucleic Acids Res* 1990;**18**:3227–31.
108 **Schwaab R**, Ludwig M, Kochhan L, *et al*. Detection and characterisation of two missense mutations at a cleavage site in the factor VIII light chain. *Thromb Res* 1991;**61**:225–34.
109 **Hamaguchi M**, Matsushita T, Tanimoto M, *et al*. Three distinct point mutations in the factor IX gene of three Japanese CRM+ hemophilia B patients (factor IX BMNagoya 2, factor IX Nagoya 3 and 4). *Thromb Haemost* 1991;**65**:514–20.
110 **Bertina RM**, van dL, I, Mannucci PM, *et al*. Mutations in hemophilia Bm occur at the Arg180–Val activation site or in the catalytic domain of factor IX. *J Biol Chem* 1990;**265**:10876–83.
111 **Noyes CM**, Griffith MJ, Roberts HR, *et al*. Identification of the molecular defect in factor IX Chapel Hill: substitution of histidine for arginine at position 145. *Proc Natl Acad Sci U S A* 1983;**80**:4200–2.
112 **Toomey JR**, Stafford D, Smith K. Factor IX Albuquerque (arginine 145 to cysteine) is cleaved slowly by factor XIa and has reduced coagulant activity. *Blood* 1988;**72**:312a.
113 **Monroe DM**, McCord DM, Huang MN, *et al*. Functional consequences of an arginine 180 to glutamine mutation in factor IX Hilo. *Blood* 1989;**73**:1540–4.
114 **Hamer RJ**, Koedam JA, Beeser-Visser NH, *et al*. Factor VIII binds to von Willebrand factor via its Mr-80,000 light chain. *Eur J Biochem* 1987;**166**:37–43.
115 **Shima M**, Scandella D, Yoshioka A, *et al*. A factor VIII neutralizing monoclonal antibody and a human inhibitor alloantibody recognizing epitopes in the C2 domain inhibit factor VIII binding to von Willebrand factor and to phosphatidylserine. *Thromb Haemost* 1993;**69**:240–6.
116 **Saenko EL**, Scandella D. The acidic region of the factor VIII light chain and the C2 domain together form the high affinity binding site for von willebrand factor. *J Biol Chem* 1997;**272**:18007–14.
117 **Jacquemin M**, Lavend'homme R, Benhida A, *et al*. A novel cause of mild/moderate hemophilia A: mutations scattered in the factor VIII C1 domain reduce factor VIII binding to von Willebrand factor. *Blood* 2000;**96**:958–65.
118 **Leyte A**, Van Schijndel HB, Niehrs C, *et al*. Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. *J Biol Chem* 1991;**266**:740–6.
119 **Higuchi M**, Wong C, Kochhan L, *et al*. Characterization of mutations in the factor VIII gene by direct sequencing of amplified genomic DNA. *Genomics* 1990;**6**:65–71.
120 **Pittman DD**, Wang JH, Kaufman RJ. Identification and functional importance of tyrosine sulfate residues within recombinant factor VIII. *Biochemistry* 1992;**31**:3315–25.
121 **Gilles JG**, Lavend'homme R, Peerlinck K, *et al*. Some factor VIII (FVIII) inhibitors recognise a FVIII epitope(s) that is present only on FVIII-vWF complexes. *Thromb Haemost* 1999;**82**:40–5.
122 **Nishino M**, Girma JP, Rothschild C, *et al*. New variant of von Willebrand disease with defective binding to factor VIII. *Blood* 1989;**74**:1591–9.
123 **Mazurier C**. von Willebrand disease masquerading as haemophilia A. *Thromb Haemost* 1992;**67**:391–6.
124 **Roelse JC**, De Laaf RT, Timmermans SM, *et al*. Intracellular accumulation of factor VIII induced by missense mutations Arg593→Cys and Asn618→Ser explains cross-reacting material-reduced haemophilia A. *Br J Haematol* 2000;**108**:241–6.
125 **Frazier D**, Smith KJ, Cheung WF, *et al*. Mapping of monoclonal antibodies to human factor IX. *Blood* 1989;**74**:971–7.
126 **Ludwig M**, Sabharwal AK, Brackmann HH, *et al*. Hemophilia B caused by five different nondeletion mutations in the protease domain of factor IX. *Blood* 1992;**79**:1225–32.
127 **Mathur A**, Zhong D, Sabharwal AK, *et al*. Interaction of factor IXa with factor VIIIa. Effects of protease domain Ca^{2+} binding site, proteolysis in the autolysis loop, phospholipid, and factor X. *J Biol Chem* 1997;**272**:23418–26.
128 **Hamaguchi N**, Bajaj SP, Smith KJ, *et al*. The role of amino-terminal residues of the heavy chain of factor IXa in the binding of its cofactor, factor VIIIa. *Blood* 1994;**84**:1837–42.
129 **Oldenburg J**, Quenzel EM, Harbrecht U, *et al*. Missense mutations at ALA-10 in the factor IX propeptide: an insignificant variant in normal life but a decisive cause of bleeding during oral anticoagulant therapy. *Br J Haematol* 1997;**98**:240–4.
130 **Jorgensen MJ**, Cantor AB, Furie BC, *et al*. Recognition site directing vitamin K-dependent gamma-carboxylation resides on the propeptide of factor IX. *Cell* 1987;**48**:185–91.
131 **Rabiet MJ**, Jorgensen MJ, Furie B, *et al*. Effect of propeptide mutations on post-translational processing of factor IX. Evidence that beta-hydroxylation and gamma-carboxylation are independent events. *J Biol Chem* 1987;**262**:14895–8.
132 **Stenflo J**, Ferlund P, Egan W, *et al*. Vitamin K dependent modifications of glutamic acid residues in prothrombin. *Proc Natl Acad Sci U S A* 1974;**71**:2730–3.
133 **Lingenfelter SE**, Berkner KL. Isolation of the human gamma-carboxylase and a gamma-carboxylase-associated protein from factor IX-expressing mammalian cells. *Biochemistry* 1996;**35**:8234–43.
134 **Willingham AK**, Matschiner JT. Changes in phylloquinone epoxidase activity related to prothrombin synthesis and microsomal clotting activity in the rat. *Biochem J* 1974;**140**:435–41.
135 **Chu K**, Wu SM, Stanley T, *et al*. A mutation in the propeptide of factor IX leads to warfarin sensitivity by a novel mechanism. *J Clin Invest* 1996;**98**:1619–25.
136 **Crossley M**, Ludwig M, Stowell KM, *et al*. Recovery from hemophilia B Leyden: an androgen-responsive element in the factor IX promoter. *Science* 1992;**257**:377–9.
137 **Briet E**, Bertina RM, van Tilburg NH, *et al*. Hemophilia B Leyden: a sex-linked hereditary disorder that improves after puberty. *N Engl J Med* 1982;**306**:788–90.
138 **Crossley M**, Brownlee GG. Disruption of a C/EBP binding site in the factor IX promoter is associated with haemophilia B. *Nature* 1990;**345**:444–6.
139 **Morgan GE**, Rowley G, Green PM, *et al*. Further evidence for the importance of an androgen response element in the factor IX promoter. *Br J Haematol* 1997;**98**:79–85.
140 **Higuchi M**, Antonarakis SE, Kasch L, *et al*. Molecular characterization of mild-to-moderate hemophilia A: detection of the mutation in 25 of 29 patients by denaturing gradient gel electrophoresis. *Proc Natl Acad Sci U S A* 1991;**88**:8307–11.
141 **Goodeve AC**, Williams I, Bray GL, *et al*. Relationship between factor VIII mutation type and inhibitor development in a cohort of previously untreated patients treated with recombinant factor VIII (Recombinate). Recombinate PUP study group. *Thromb Haemost* 2000;**83**:844–8.

142 **Vidal F**, Elisenda F, Carme A, *et al*. Rapid hemophilia A molecular diagnosis by a simple DNA sequencing procedure: identification of 14 novel mutations. *Thromb Haemost* 2001;**85**:580–3.
143 **Naylor JA**, Green PM, Montandon AJ, *et al*. Detection of three novel mutations in two haemophilia A patients by rapid screening of whole essential region of factor VIII gene. *Lancet* 1991;**337**:635–9.
144 **Montandon AJ**, Green PM, Giannelli F, *et al*. Direct detection of point mutations by mismatch analysis: application to haemophilia B. *Nucleic Acids Res* 1989;**17**:3347–58.
145 **Green PM**, Montandon AJ, Ljung R, *et al*. Haemophilia B mutations in a complete Swedish population sample: a test of new strategy for the genetic counselling of diseases with high mutational heterogeneity. *Br J Haematol* 1991;**78**:390–7.
146 **Saad S**, Rowley G, Tagliavacca L, *et al*. First report on UK database of haemophilia B mutations and pedigrees. UK haemophilia centres. *Thromb Haemost* 1994;**71**:563–70.
147 **Waseem NH**, Bagnall R, Green PM, *et al*. Start of UK confidential haemophilia A database: analysis of 142 patients by solid phase fluorescent chemical cleavage of mismatch. Haemophilia centres. *Thromb Haemost* 1999;**81**:900–5.
148 **Rowley G**, Saad S, Giannelli F, *et al*. Ultrarapid mutation detection by multiplex, solid-phase chemical cleavage. *Genomics* 1995;**30**:574–82.
149 **Costa JM**, Ernault P, Vidaud D, *et al*. Fast and efficient mutation detection method using multiplex PCR and cycle sequencing—application to haemophilia B. *Thromb Haemost* 2000;**83**:244–7.
150 **Giannelli F**, Green PM, Naylor JA. A genetic view on the etiology of the inhibitor complication. *Blood* 1996;**87**:2612.
151 **Tuddenham EG**, McVey JH. The genetic basis of inhibitor development in haemophilia A. *Haemophilia* 1998;**4**:543–5.
152 **Schwaab R**, Brackmann HH, Meyer C, *et al*. Haemophilia A: mutation type determines risk of inhibitor formation. *Thromb Haemost* 1995;**74**:1402–6.
153 **Tuddenham EG**. Molecular biological aspects of inhibitor development. *Vox Sang* 1999;**77**(suppl 1):13–16.
154 **Vianello F**, Radossi P, Tison T, *et al*. Prevalence of anti-FVIII antibodies in severe haemophilia A patients with inversion of intron 22. *Br J Haematol* 1997;**97**:807–9.
155 **Hay CR**, Ollier W, Pepper L, *et al*. HLA class II profile: a weak determinant of factor VIII inhibitor development in severe haemophilia A. UKHCDO inhibitor working party. *Thromb Haemost* 1997;**77**:234–7.
156 **Oldenburg J**, Picard JK, Schwaab R, *et al*. HLA genotype of patients with severe haemophilia A due to intron 22 inversion with and without inhibitors of factor VIII. *Thromb Haemost* 1997;**77**:238–42.
157 **Oldenburg J**, Schroder J, Schmitt C, *et al*. Small deletion/insertion mutations within poly-A runs of the factor VIII gene mitigate the severe haemophilia A phenotype. *Thromb Haemost* 1998;**79**:452–3.
158 **Young M**, Inaba H, Hoyer LW, *et al*. 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**:565–73.
159 **Lenting PJ**, Neels JG, van den Berg BM, *et al*. The light chain of factor VIII comprises a binding site for low density lipoprotein receptor-related protein. *J Biol Chem* 1999;**274**:23734–9.
160 **Saenko EL**, Yakhyaev AV, Mikhailenko I, *et al*. Role of the low density lipoprotein-related protein receptor in mediation of factor VIII catabolism. *J Biol Chem* 1999;**274**:37685–92.
161 **Schwarz HP**, Lenting PJ, Binder B, *et al*. Involvement of low-density lipoprotein receptor-related protein (LRP) in the clearance of factor VIII in von Willebrand factor-deficient mice. *Blood* 2000;**95**:1703–8.
162 **Neels JG**, van den Berg BM, Mertens K, *et al*. Activation of factor IX zymogen results in exposure of a binding site for low-density lipoprotein receptor-related protein. *Blood* 2000;**96**:3459–65.

Reference linking to full text of more than 200 journals

Toll free links

You can access the FULL TEXT of articles cited in the *Journal of Clinical Pathology* online if the citation is to one of the more than 200 journals hosted by HighWire (http://highwire.stanford.edu) without a subscription to that journal. There are also direct links from references to the Medline abstract for other titles.

www.jclinpath.com