MspI polymorphic site within the Factor IX gene

Localization of the site and an improved method for detection

Debra L. Freedenberg[1,3], Shi-Han Chen[1,3], Kotoku Kurachi[2*], and C. Ronald Scott[1,3]

[1] Department of Pediatrics and [2] Department of Biochemistry, University of Washington, Seattle, WA 98195, USA
[3] Children's Hospital and Medical Center, Seattle, WA 98105, USA

Summary. We have localized the position of the MspI polymorphic site that exists within the factor IX gene. The location of the MspI polymorphic site is within intron D, 1.9 kb upstream from the beginning of exon V and 4 kb downstream from a known polymorphic TaqI site. The use of a specific genomic probe simplifies the interpretation of the MspI polymorphism by reducing the number of non-overlaping DNA fragments to three bands; 2.4, 3.4, and 5.8 kb.

Fig. 1. Location of polymorphic restriction sites in the factor IX gene and the location of the genomic probe used for localization of the MspI Site. Exons I–VIII of the factor IX gene are represented by rectangular blocks labeled I–VIII. The intervening areas are introns. The *upper solid bar* represents the genomic probe (the sequence between EcoRI-6 and EcoRI-8 sites) used for localization of the MspI site. The polymorphic site of TaqI, DdeI, and XmnI are indicated below the *solid line*. * FIXOPl is a polymorphism detected by oligonucleotide probes (Winship and Brownlee 1986)

Introduction

Hemophilia B has a frequency of 1/30,000 in newborn males. Molecular cloning of factor IX cDNA (Anson et al. 1984; Choo et al. 1982; Jaye et al. 1983; Kurachi and Davies 1982) has allowed the detection of several restriction fragment length polymorphisms (RFLPs); TaqI, XmnI, DdeI, and BamHI (Camerino et al. 1984; Giannelli et al. 1984; Winship et al. 1984; Hay et al. 1986). These polymorphisms are useful for prenatal diagnosis and carrier detection of the gene for hemophilia B. A recently detected RFLP in human DNA digested with the endonuclease MspI (Camarino et al. 1985) is manifest as a variable restriction pattern of either 2.4 kb or 5.8 kb fragments on Southern blots when hybridized with factor IX cDNA. The genomic location of this polymorphic site remained elusive but was believed to be outside of exons II, III, and IV and their neighboring introns (Camarino et al. 1985). We have utilized a specific genomic probe which improves detection of the 2.4 kb fragment and have localized the MspI polymorphic site to intron D.

Materials and methods

DNA was prepared from normal individuals and individuals with hemophilia B from approximately 10 ml anticoagulated blood as previously described (Murray et al. 1983). Restriction enzyme digestions were followed by 0.8% agarose gel electrophoresis and Southern blot transfer onto nitrocellulose filters as described by Southern (1975). The DNA was hybridized with a factor IX cDNA probe or a genomic probe labeled with [α ³²P]-deoxynucleotides. Filters were then washed and
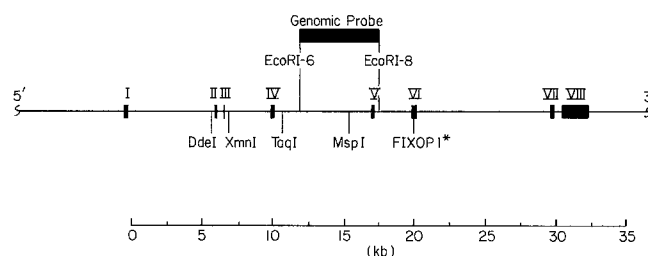
exposed to Kodak AR film between a pair of intensifying screens for a period of 3–10 days.

The nitrocellulose filters were washed in 0.1% sodium dodecyl sulfate (SDS) at 70°C for one hour after development of the autoradiogram. These filters were available for re-hybridization with the probe that had not been used initially (either a genomic probe or the cDNA probe).

The cDNA probe and the genomic probe were labeled with [α ³²P]-deoxynucleotides by a nick translation reaction method to 1–3 × 10⁸ cpm/μg DNA. The cDNA probe includes the sequence coding for the signal peptide, mature protein, and 3′ noncoding region. The genomic probe is approximately 5.5 kb in length and contains the sequence starting at the 6th EcoRI site to the 8th EcoRI site (Fig. 1). The probe contains all of exon V and most of intron D.

Results and discussion

When human DNA digested with MspI is hybridized with the cDNA probe there are five constant bands (1.4 kb, 1.6 kb, 2.5 kb, 4.3 kb, 5.5 kb) and two variable bands (5.8 kb and 2.4 kb) as shown in Fig. 2. Using the genomic probe with DNA obtained from males, we find there is either a single 5.8 kb band or a 3.4 and 2.4 kb band. Females may have all three bands, or may be homozygous for either (Fig. 3). All constant fragments present in the cDNA hybridizations are not detect-
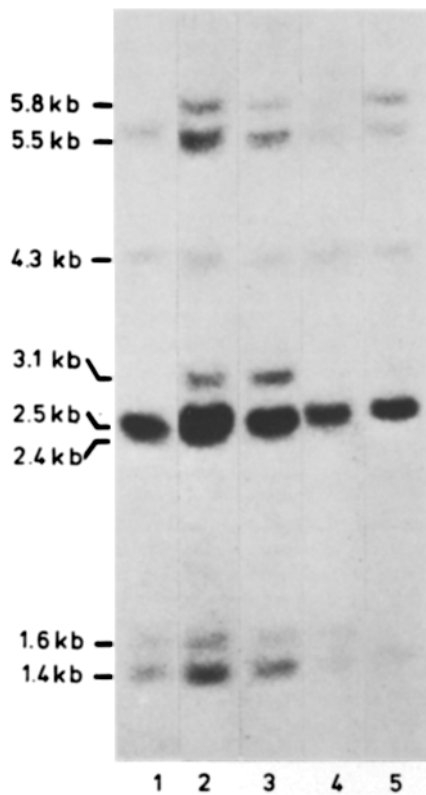
Fig.2. MspI fragments after Southern blots and hybridization with a cDNA probe. *Lane 1* is a male with the 2.4 fragment. *Lanes 2* and *3* are females who have both the 2.4 and 5.8 fragments. *Lane 4* is a female homozygous for the 2.4 fragment. *Lane 5* is DNA from a male with the 5.8 fragment. Note the difficulty in distinguishing the variable 2.4 kb fragment from the large constant 2.5 kb band. The 3.1 kb fragment is a band that is occasionally present and is not inherited in a Mendelian manner. It most likely represents incomplete digestion by MspI

Fig.3. MspI restriction fragments after Southern blot and hybridization with the genomic probe. *Lane 1* is a hemizygous male with the 5.8 fragment. *Lane 2* is a hemizygous male with the 2.4 fragment. *Lanes 3* and *4* represent heterozygous females in whom the 2.4 and 5.8 fragments are present. *Lane 5* is a female who is homozygous for the 2.4 fragment. The 3.4 kb fragment exists only in the presence of the 2.4 kb fragment

Table 1. Linkage of TaqI/MspI haplotypes. MspI and TaqI haplotypes determined from 51 informative chromosomes in a population of unrelated caucasions (9 males, 21 females). The MspI and TaqI restriction sites are in apparent linkage disequilibrium ($P < 0.01$)

Haplotype	Observed	Expected
MspI 2.4 / TaqI 1.8	34	27.5
MspI 2.4 / TaqI 1.3	5	11.2
MspI 5.8 / TaqI 1.8	2	8.7
MspI 5.8 / TaqI 1.3	10	3.6

able. Thus, the interpretation of the MspI polymorphic patterns has been simplified by using the genomic probe.

The presence of the MspI polymorphism when human DNA is hybridized to the genomic probe that contains only the nucleotide sequence from exon V and intron D, indicates that the polymorphic site must be near this region. From the published nucleotide sequence (Yoshitake et al. 1985), there exist MspI sites that are close to the EcoRI-6 and EcoRI-8 sites of the factor IX gene. The distance between these MspI sites is approximately 6.0 kb. This agrees with the size of the fragment that is detected by the genomic probe when the polymorphic MspI site is absent: 5.8 kb. When the MspI polymorphic site is present, we detect 2.4 kb and 3.4 kb fragments. This indicates that the polymorphic site lies between the two constant MspI sites that border the EcoRI-6 and the EcoRI-8 sites. The absence of a MspI recognition sequence between these two sites in the published sequence data (Yoshitake et al. 1985) may represent the sequencing of the factor IX gene from an individual who lacks the MspI polymorphic site and has the 5.8 kb fragment. There are at least ten candidates for the polymorphic site that differ by only one base from the MspI recognition sequence. These sites are within several hundred base pairs of each other. The intragenic position of these sites is approximately 1.9 kb upstream from the beginning of exon V and 4.5 kb downstream from the TaqI polymorphic site in intron D (Fig.1). Hence, we have established
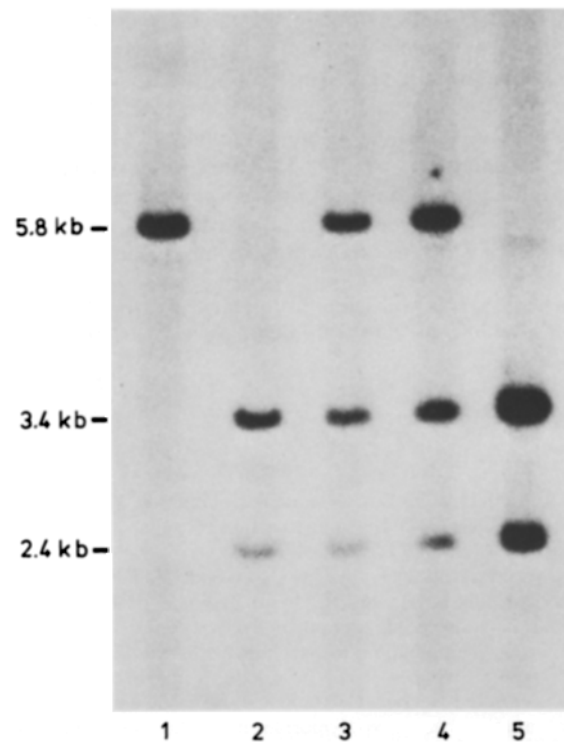
the location of the polymorphic site based on the use of our genomic probe and the published sequence data of the factor IX gene. Camarino et al. (1985) had originally estimated the polymorphic site to be within 10 kb of the variable TaqI site in intron D. This is in agreement with our findings.

In 51 unrelated Caucasian chromosomes we found that the 2.4 kb and 5.8 kb fragments have a frequency of 0.76 and 0.24 respectively. Combining these data with other published information (Camarino et al. 1985), the total number of chromosomes evaluated for the MspI polymorphism is 125. The combined data give a frequency of 0.22 for the 5.8 kb allele and 0.78 for the 2.4 kb allele in Caucasian populations (95% confidence intervals are 0.14 to 0.29 and 0.70 to 0.85, respectively).

The MspI site continues to show linkage disequilibrium with the variable TaqI site ($P < 0.01$; Table 1). The TaqI 1.8/ MspI 2.4 haplotype is the most common in our population.

Elucidation of the location of the MspI polymorphic site within intron D indicates that it will be of value in studying families with hemophilia B. It is approximately in the middle of the gene and has a polymorphic frequency that is useful for prenatal diagnosis or carrier detection. Although the MspI site is in linkage disequilibrium with the TaqI site, we have encountered a family in which the TaqI polymorphism was uninformative while the MspI polymorphism provided the necessary information with which to offer genetic counseling.

References

Anson DS, Choo KH, Rees DJG, Giannelli E, Gould K, Huddleston JA, Brownlee GG (1984) The gene structure of human antihaemophilic factor IX. EMBO J 3:1053–1060

Camerino GG, Grzeschik KH, Jaye M, De La Salle H, Tolstoshev P, Lecocq JP, Heilig R, Mandel JL (1984) Regional localization on the human X chromosome and polymorphism of the coagulation factor IX gene (hemophilia B locus). Proc Natl Acad Sci USA 81:498–502

Camerino G, Oberle I, Drayna D, Mandel JL (1985) A new MspI restriction fragment length polymorphism in the hemophilia B locus. Hum Genet 71:79–81

Choo KH, Gould KG, Rees DJG, Brownlee GG (1982) Molecular cloning of the gene for human anti-hemophilic factor IX. Nature 299:178–180

Gianelli F, Choo KH, Winship PR, Rizza CR, Anson DS, Rees DJG, Ferrari N, Brownlee GG (1984) Characterization and use of an intragenic polymorphic marker for detection of carriers of hemophilia B (factor IX deficiency). Lancet I:239–241

Hay CW, Robertson KA, Yong SL, Thompson AR, Growe GH, MacGillivray RTA (1986) Use of a BamHI polymorphism in the factor IX gene for the determination of hemophilia B carrier status. Blood 67:1508–1511

Jaye M, De La Salle H, Schamber F, Balland A, Kohli V, Findeli A, Tolstoshev P, Lecocq JP (1983) Isolation of a human antihaemophilic factor IX cDNA clone using a unique 52 base synthetic oligonucleotide probe deduced from the amino acid sequence of bovine factor IX. Nucleic Acids Res 11:2325–2335

Kurachi K, Davies EW (1982) Isolation and characterization of a cDNA coding for human factor IX. Proc Natl Acad Sci USA 78:6461–6464

Murray JC, Demopulos CM, Lawn RM, Motulsky AG (1983) Molecular genetics of human serum albumin: restriction enzyme fragment length polymorphism and analbuminemia. Proc Natl Acad Sci USA 80:5951–5955

Southern EN (1975) Detection of specific sequences among DNA fragment separated by gel electrophoresis. J Mol Biol 98:503–517

Winship RR, Brownlee GG (1986) Diagnosis of haemophilia B carriers using intragenic oligonucleotide probes. Lancet II:218–219

Winship RR, Anson DS, Rizza CR, Brownlee GG (1984) Carrier detection in haemophilia B using two further intragenic restriction fragment length polymorphisms. Nucleic Acids Res 12:8861–8872

Yoshitake S, Schach BG, Foster DC, Davie EW, Kurachi K (1985) Nucleotide sequence of the gene for human factor IX (antihemophilic factor B). Biochemistry 24:3736–3750