



Combined analysis of the *MspI* and *XbaI* polymorphisms in intron 22 of the factor VIII gene for detection of hemophilia A in a Korean population

S.H. Park¹, N. Chung¹, M.R. Lee², S.K. Yoo² and Y.M. Choi³

¹Division of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul, Korea

²Korea Hemophilia Foundation Clinic, Seoul, Korea

³Department of Obstetrics and Gynecology, The Institute of Reproductive Medicine and Population, Medical Research Center, Seoul National University College of Medicine, Seoul, Korea

Corresponding author: Y.M. Choi
E-mail: ymchoi@snu.ac.kr

Genet. Mol. Res. 11 (1): 1-9 (2012)

Received June 6, 2011

Accepted September 5, 2011

Published January 9, 2012

DOI <http://dx.doi.org/10.4238/2012.January.9.1>

ABSTRACT. To determine the usefulness of *MspI*/int22h-1 (intron 22 homologous region-1) polymorphism of the factor VIII gene for molecular genetic diagnosis of hemophilia A in the Korean population, *MspI*/intron 22 and *XbaI*/intron 22 polymorphisms were analyzed in 101 unrelated Korean families with severe hemophilia A. The expected heterozygosity rates of *MspI*/int22h-1 and *XbaI*/int22h-1 polymorphisms were 49.5 and 43.6%, respectively; these polymorphisms were not in complete linkage disequilibrium. Combined analysis using both polymorphisms provided an informative rate of 66.3%. These results suggest that PCR analysis of the *MspI*/int22h-1 polymorphism of the factor VIII gene would be useful for

carrier detection and prenatal diagnosis of hemophilia A in the Korean population.

Key words: Factor VIII gene; Hemophilia A; Intron 22; *MspI*; *XbaI*; Molecular genetic diagnosis

INTRODUCTION

Hemophilia A is one of the most common bleeding disorders in humans, affecting 1 in 5,000-10,000 males. The disorder is caused by mutations in the factor VIII gene (Gitschier et al., 1984; Bowen, 2002). The factor VIII gene is located toward the end of the long arm of the X-chromosome (Xq28) and, at 180 kb and containing 26 exons, is an extremely large gene (Gitschier et al., 1984). More than 900 different kinds of mutations are known to cause hemophilia A (Tuddenham et al., 1994), but except for one case of inversion, no specific mutational hot spot has been reported. The inversion involves intron 22 of the factor VIII gene and is found in 40-50% of patients with severe hemophilia A (Lakich et al., 1993; Naylor et al., 1993). The inversion can be identified by a direct analysis of the mutation using a long PCR assay method (Liu et al., 1998). However, except for this inversion, it is very difficult to identify specific gene mutations in patients in order to detect carriers and perform prenatal diagnoses. Thus, in cases lacking the inversion, gene polymorphisms have been commonly used for mutational linkage analysis to diagnose the disease.

To perform molecular genetic diagnoses of hemophilia A, gene polymorphisms are most commonly analyzed with extragenic markers such as St14 VNTR (variable number of tandem repeats) (Richards et al., 1991), intragenic markers such as the *BclII*/intron 18 polymorphism (Gitschier et al., 1985), hypervariable dinucleotide repeats of intron 13 (Lalloz et al., 1991) and intron 22 (Lalloz et al., 1992, 1994), and the *XbaI*/intron 22 polymorphism (Wion et al., 1986). The intragenic intron 22 (int22h-1; intron 22 homologous region-1) of the factor VIII gene has a similar gene sequence at two locations (int22h-2, int22h-3) of the extragenic Xq telomere, where gene polymorphisms also exist (Levinson et al., 1990, 1992; Naylor et al., 1995). Therefore, specific information about int22h-1 gene polymorphisms would be important for accurate diagnosis. The *XbaI* and *MspI* polymorphisms are located within a 9.5-kb stretch of the int22h-1 sequence. The extragenic homologous sequences, int22h-2 and int22h-3, also have polymorphic sites for *XbaI* and *MspI* (Chan et al., 1989; Naylor et al., 1995; Bowen et al., 2000). Thus, detection of *XbaI*/int22h-1 and *MspI*/int22h-1 polymorphisms at the three polymorphic sites of int22h would require efficient amplification of those specific sequences (De Brasi et al., 1999; El-Maarri et al., 1999).

The *MspI* polymorphism is located 737 bases downstream from the *XbaI* polymorphism (Bowen et al., 2000). Despite their close proximity to each other, these polymorphisms are known to not be in complete linkage disequilibrium. Bowen et al. (2000) reported that when 85 male residents of South Wales were studied, the informative rate increased to 60% through the combined use of *MspI* and *XbaI* polymorphisms of int22h-1. This finding, however, has not been shown to be applicable to people of other regions, especially Asian countries. With this thought in mind, we investigated the degree of linkage disequilibrium of both *XbaI* and *MspI* polymorphisms in a Korean population, which could be used for carrier detection and prenatal diagnosis of hemophilia A.

MATERIAL AND METHODS

Subjects

We recruited 101 unrelated Korean families with hemophilia A during the period 1992 to 2004, who visited Korea Hemophilia Foundation for the purpose of carrier detection and prenatal diagnosis of hemophilia A. The review board for human research of Seoul National University Hospital approved this project.

Genomic DNA extraction

Genomic DNA was extracted from peripheral blood with a Wizard genomic DNA purification kit (Promega, Madison, WI, USA). The concentration of extracted DNA was measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Each DNA sample was analyzed after normalizing the concentration to 50 ng/ μ L.

Analysis of *XbaI*/int22h-1 polymorphism

PCR was performed in a 25- μ L total reaction volume containing 200 ng genomic DNA, Takara LA *Taq* buffer (Takara Bio Inc., Kyoto, Japan), 7.5% DMSO (dimethyl sulfide), 250 μ M each of dGTP and 7-deaza-dGTP (Boehringer Mannheim, Mannheim, Germany), 500 μ M of the other dNTPs (Boehringer Mannheim), 1 U LA *Taq* DNA polymerase (Takara Bio Inc.), 0.2 μ M each of the forward and reverse primers. The forward primer was 5'-GCCCTGCCGTGCCATTACTGATGACATTATGCTGAC-3' and the reverse primer was 5'-TTCAACACGACCACCATCTCTCAAGTGGCC-3'. PCR cycling was performed in an MJR PTC200 thermocycler (MJ Research, Waltham, MA, USA) as follows: 94°C for 2 min, followed by 30 cycles at 94°C for 12 s, 65°C for 30 s, 60°C for 7 min, with auto-extension of 20 s for the last 20 cycles, followed by 15 min at 68°C. Approximately 5 μ L of the resulting PCR product was run on a 1% agarose gel to check for the presence of the desired product. Digestion was performed in a 12- μ L total reaction volume containing 10 μ L PCR product, 10 U *XbaI* enzyme (Promega), 1.2 μ L *XbaI* restriction buffer (buffer D), and bovine serum albumin to a final concentration of 0.1 mg/mL. The digestion reaction was kept at 37°C overnight. The digestion product was run on a 1% agarose gel at 100 V until optimum band separation was achieved. The gel was stained with ethidium bromide and the bands were visualized by exposure to UV light.

Analysis of *MspI*/int22h-1 polymorphism

Long PCR was performed as reported by De Brasi et al. (1999). Nested PCR was performed using the long PCR product as substrate (1 μ L of a 1:500 dilution). PCR was performed in a 25- μ L total reaction volume containing 1 μ L diluted long PCR product, PCR buffer (Takara *Taq* buffer), 500 μ M dNTP mixture, 1 U *Taq* DNA polymerase (Takara Bio Inc.), 0.4 μ M each of the forward and reverse primers. The forward primer was 5'-GGTGCTCAGTAGCCTGTCGTTGTG-3' and the reverse primer was 5'-GCCACTACGCTCAGGTCCTGAGTC-3'. PCR cycling was performed in an MJR PTC200 thermocycler (MJ Research), as

follows: 94°C for 5 min, followed by 30 cycles at 94°C for 30 s, 58°C for 60 s, 72°C for 60 s. Approximately 5 µL of the resulting PCR product was run on a 1% agarose gel to check for the presence of the desired product. Digestion was performed in a 12-µL total reaction volume containing 10 µL PCR product, 10 U *MspI* (New England Biolabs Inc., Beverly, MA, USA) and 1.2 µL NEB Buffer 2. The digestion reaction was kept at 37°C. This digestion product was run on a 3.5% agarose gel at 100 V until optimum band separation was achieved. The gel was stained with ethidium bromide and the bands were visualized by exposure to UV light.

Statistical analysis

To determine whether each single nucleotide polymorphism site was in Hardy-Weinberg equilibrium (HWE), the distributions of observed genotype frequencies and expected genotype frequencies were compared by the chi-square test using a website (<http://ihg2.helmholtz-muenchen.de/cgi-bin/hw/hwa1.pl>). Haplotypes were estimated with Haploview (ver. 4.0). Linkage disequilibrium (LD) between the two polymorphic sites was estimated as relative disequilibrium (D') and correlation factor (r^2).

RESULTS

One hundred and one unrelated patients with severe hemophilia A, as well as their family members, were analyzed for the *XbaI*/int22h-1 and *MspI*/int22h-1 polymorphisms. Figure 1 shows schematically the positions of the polymorphic sites and amplification products for the analysis of *XbaI*/int22h-1 and *MspI*/int22h-1 polymorphisms.

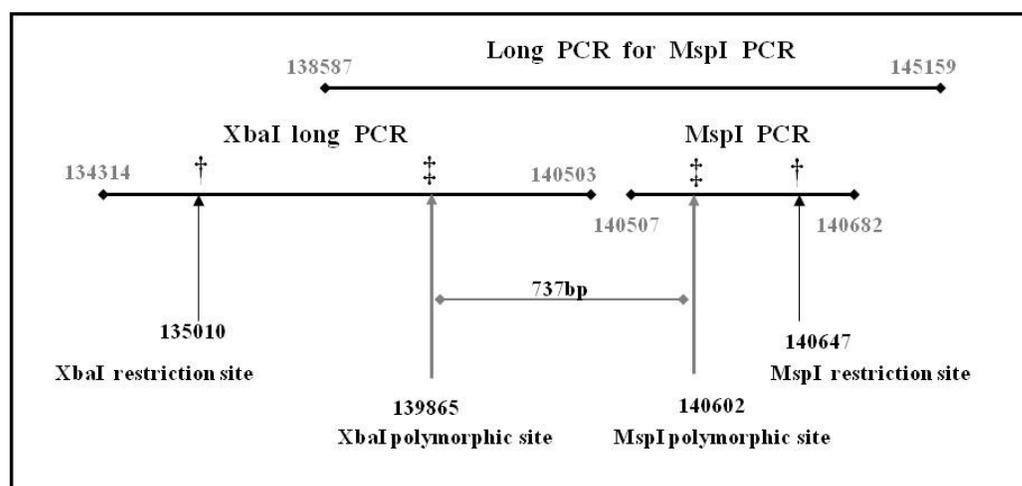
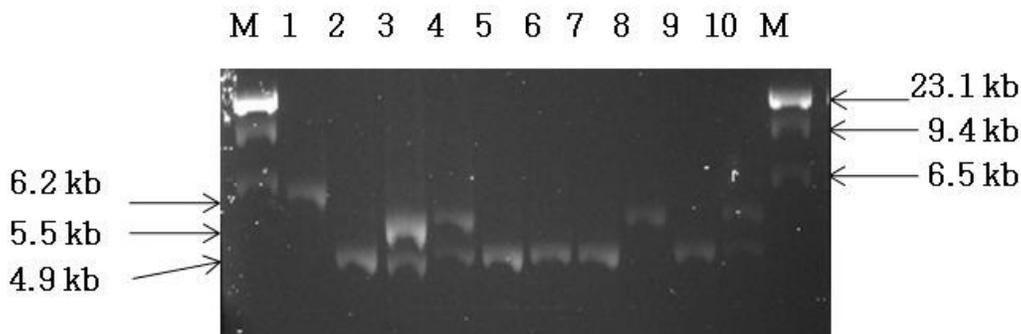


Figure 1. Diagram of the intron 22 region for analysis of *XbaI* and *MspI* polymorphisms. This diagram shows the positions of restriction sites and PCR amplifications based on a published sequence (GB AY769950, 192029 bp). † indicates the internal control sites of each restriction enzyme and ‡ indicates the polymorphic target sites of each restriction enzyme.

***XbaI*/int22h-1 polymorphism**

The amplification product of *XbaI* long PCR was 6.2 kb (Figure 2A, lane 2). The PCR product has an invariant 696-bp *XbaI* site located at the 5' end of the PCR product, which provided an internal digestion control (data not shown). While 5.5- and 0.7-kb fragments were obtained for *XbaI* A (-) DNA, the 5.5-kb fragment was further digested into 4.9- and 0.6-kb fragments for *XbaI* A (+) DNA (Figure 2A). When the patients' mothers' DNA sequences were analyzed for the *XbaI*/int22h-1 polymorphism, the genotype (+,+) was found in 43 of 101 subjects, genotype (-,-) in 14 of 101, and genotype (+,-) in 44 of 101 (Table 1). Thus, the heterozygosity rate was determined to be 43.6%. The (+) allele had a frequency of 0.644 and the (-) allele had a frequency of 0.356 (allele +/- = 0.644/0.356 ± 0.0345, 95%CI).

A



B

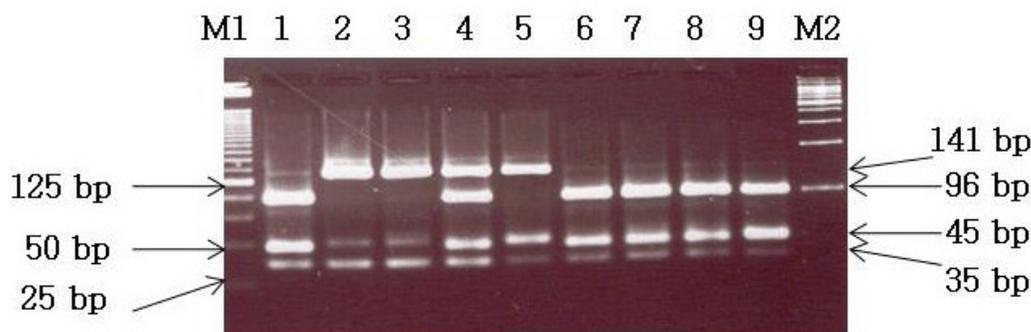


Figure 2. Analysis of *XbaI* and *MspI* RFLPs. **A.** Agarose gel electrophoresis of the long PCR product and *XbaI* digestion products. Lane M = λ /HindIII marker; lane 1 = undigested product; lanes 3, 4 = (+,-); lanes 2, 5, 6, 7, 9 = (+,+); lane 8 = (-,-); lane 10 = negative control; left side = the product size and right side = the λ /HindIII marker size. **B.** Agarose gel electrophoresis of *MspI* digestion products after nested PCR. Lane M1 = 25-bp marker; lanes 1, 6, 7, 8, 9 = (+,+); lane 4 = (+,-); lanes 2, 3, 5 = (-,-); lane M2 = 100-bp marker; left side = the 25-bp marker size and right side = the product size.

Table 1. Frequencies of *XbaI* and *MspI* genotypes.

Genotype	Frequency	
	Number	% of total
<i>XbaI</i>		
-,-	14	13.9
+,-	44	43.6
+,+	43	42.6
<i>MspI</i>		
-,-	14	13.9
+,-	50	49.5
+,+	37	36.6

MspI/int22h-1 polymorphism

For polymorphism analysis at the *MspI* polymorphic site, the long PCR product was used as a template for nested PCR. The long PCR product size was 6.6 kb and the nested PCR product size was 176 bp. As in the *XbaI* polymorphism analysis, the invariant 35-bp *MspI* site, located at the 3' end of the PCR product, provided an internal digestion control. While fragments of 141 and 35 bp were obtained for *MspI* A (-) DNA, the 141-bp fragment was further digested to fragments of 96 and 45 bp for *MspI* A (+) DNA (Figure 2). When the patients' mothers' DNA sequences were analyzed for the *MspI/int22h-1* polymorphism, the genotype (+,+) was found in 37 of 101 subjects, genotype (-,-) in 14 of 101, and genotype (+,-) in 50 of 101 (Table 1). Thus, the heterozygosity rate was determined to be 49.5%. The (+) allele had a frequency of 0.61 and the (-) allele had a frequency of 0.39 (allele +/- = 0.61/0.39 ± 0.0335, 95%CI).

Linkage analysis between *MspI/int22h-1* and *XbaI/int22h-1*

When the patients' mothers' DNA sequences were analyzed for *MspI* and *XbaI* haplotypes, the two polymorphic sites had a correlation factor of $r^2 = 0.022$, which indicates a lack of complete linkage disequilibrium. This finding is surprising, because the distance between these polymorphic sites is only 737 bp (Figure 1). Therefore, when *MspI* and *XbaI* sites were analyzed together using allelic frequency and linkage data, a more informative rate was obtained. When the *XbaI* site was homozygous and therefore uninformative, additional informative rates of 16.8 and 5.9% were obtained for homozygous females with genotypes of *XbaI* (+,+) and *XbaI* (-,-), respectively. While analysis using only the *XbaI* site gave a 43.6% informative rate, analysis of the *XbaI* and *MspI* sites together gave an additional 22.7% informative rate (Table 2). Thus, the informative rate increased to 66.3% with the use of both polymorphic sites.

Table 2. Frequencies of *XbaI* and *MspI* haplotypes.

Genotype		Frequency	
<i>XbaI</i>	<i>MspI</i>	Number	% of total
-,-	-,-	8	7.9
-,-	+,-	6	5.9
-,-	+,+	0	0
+,-	-,-	5	5.0
+,-	+,-	27	26.7
+,-	+,+	12	11.9
+,+	-,-	1	1.0
+,+	+,-	17	16.8
+,+	+,+	25	24.8
Total		101	100

DISCUSSION

It is known that the polymorphic homologous regions int22h-2 and int22h-3 exist outside of intron 22 of the factor VIII gene on the X-chromosome (Naylor et al., 1995). Thus, detection of the extragenic homologous regions could result in an incorrect diagnosis during polymorphic analysis of intron 22. Southern blot analyses of intron 22 polymorphisms could also result in an incorrect diagnosis due to the presence of two adjacent bands (Van de Water et al., 1994). Additionally, when conventional PCR was used, the 3' regions were amplified together and polymorphisms of the int22h-1 region could not be accurately distinguished (Kogan et al., 1987). Because of these difficulties, the use of the *Xba*I/int22h-1 polymorphism has been restricted to the clinical diagnosis of hemophilia A. However, El-Maarri et al. (1999) developed a PCR-based analysis method that is specific for the *Xba*I/int22h-1 polymorphism. When El-Maarri et al. (1999) used this method to analyze the DNA of Turkish patients, they reported that *Xba*I/int22h-1 polymorphism analysis was effective for 60% of families who were uninformative with other intragenic markers. In subsequent studies, *Xba*I/int22h-1 polymorphism was reported to have a heterozygosity rate of 48-50%, which did not vary for different ethnic groups (Wion et al., 1986; Chan et al., 1989; Peake et al., 1993).

In the present study, the expected heterozygosity rate for the *Xba*I/int22h-1 polymorphism in the Korean population was 43.6%, a value similar to the 48% heterozygosity rate in the Caucasian population (El-Maarri et al., 1999). The value for the *Xba*I/int22h-1 polymorphism in the Korean population was much higher than the value (20%) for the *Bcl*II/intron 18 polymorphism in this same population (Choi et al., 2000). Therefore, the *Xba*I/int22h-1 polymorphism is more useful for molecular genetic diagnosis of hemophilia A in the Korean population. In addition, the *Msp*I polymorphism of int22h-1, which is 737 bp away from the *Xba*I polymorphism site, had a heterozygosity rate of 49.5%, and allelic frequencies were 0.61 and 0.39, respectively, for the (+) and (-) alleles (allele +/- = 0.61/0.39 ± 0.0335, 95%CI) in this study. These results are similar to those of Bowen et al. (2000), where the heterozygosity rate was 0.46, and allelic frequencies were 0.65 and 0.35, respectively, for the (+) and (-) alleles in 85 male hemophilia A patients in South Wales. This information suggested that a slightly higher informative rate was obtained in the Korean population.

Even though the informative rate obtained with either the *Xba*I or *Msp*I polymorphism of int22h-1 is slightly lower than those obtained with factor VIII introns 13 and 22 microsatellites (67 and 34%, respectively), the *Xba*I or *Msp*I polymorphisms could compensate for weak points of the microsatellites, the allelic frequencies of which could fluctuate between generations (Jeffreys et al., 1988; Rubinsztein et al., 1995; Lee et al., 2009). Thus, we could obtain better diagnoses if introns 13 and 22 microsatellite data were first verified using the *Xba*I or *Msp*I polymorphisms.

Our present study also showed that no complete LD existed between the *Msp*I polymorphism and *Xba*I polymorphism in a Korean population and that simultaneous use of both polymorphisms provided an informative rate of 66.3%, indicating a high quality of molecular markers for hemophilia A. However, the experimental procedure for detection of the *Msp*I polymorphism includes long PCR and nested PCR to exclude extragenic int22h-2 and int22h-3. The steps of this procedure are complicated and samples could be contaminated, resulting in an erroneous interpretation. In addition, DNA integrity has been the most important factor in similar experiments. However, if an experienced researcher performs the experiments

and DNA integrity is guaranteed, a much higher informative rate can be obtained using the *MspI* polymorphism of int22h-1 for carrier detection and prenatal diagnosis of hemophilia A.

ACKNOWLEDGMENTS

Research supported by the Korea Hemophilia Foundation in Korea (2005).

REFERENCES

- Bowen DJ (2002). Haemophilia A and haemophilia B: molecular insights. *Mol. Pathol.* 55: 1-18.
- Bowen DJ, De Brasi CD, Larripa IB and Collins PW (2000). 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.* 111: 544-548.
- Chan V, Tong TM, Chan TP, Tang M, et al. (1989). Multiple XbaI polymorphisms for carrier detection and prenatal diagnosis of haemophilia A. *Br. J. Haematol.* 73: 497-500.
- Choi YM, Hwang D, Choe J, Jun JK, et al. (2000). Carrier detection and prenatal diagnosis of hemophilia A in a Korean population by PCR-based analysis of the BclII/intron 18 and St14 VNTR polymorphisms. *J. Hum. Genet.* 45: 218-223.
- De Brasi CD, Bowen DJ, Collins PW and Larripa IB (1999). Specific analysis of the intron 22 XbaI polymorphism of the human factor VIII gene using long-distance PCR. *Br. J. Haematol.* 107: 566-568.
- El-Maarri O, Oldenburg J and Caglayan SH (1999). Intron 22-specific long PCR for the Xba I polymorphism in the factor VIII gene. *Br. J. Haematol.* 105: 1120-1122.
- Gitschier J, Wood WI, Goralka TM, Wion KL, et al. (1984). Characterization of the human factor VIII gene. *Nature* 312: 326-330.
- Gitschier J, Drayna D, Tuddenham EGD, White RL, et al. (1985). Genetic mapping and diagnosis of haemophilia A achieved through a BclI polymorphism in the factor VIII gene. *Nature* 314: 738-740.
- Jeffreys AJ, Royle NJ, Wilson V and Wong Z (1988). Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* 332: 278-281.
- Kogan SC, Doherty M and Gitschier J (1987). An improved method for prenatal diagnosis of genetic diseases by analysis of amplified DNA sequences. Application to hemophilia A. *N. Engl. J. Med.* 317: 985-990.
- Lakich D, Kazazian HH Jr, Antonarakis SE and Gitschier J (1993). Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nature Genet.* 5: 236-241.
- Lalloz MR, McVey JH, Pattinson JK and Tuddenham EG (1991). Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet* 338: 207-211.
- Lalloz MRA, McVey JH, Michaelides K and Tuddenham EG (1992). Haemophilia A diagnosis by analysis of a novel dinucleotide tandem repeat sequence within the factor VIII gene. *Br. J. Haematol.* 80: 3a.
- Lalloz MR, Schwaab R, McVey JH, Michaelides K, et al. (1994). Haemophilia A diagnosis by simultaneous analysis of two variable dinucleotide tandem repeats within the factor VIII gene. *Br. J. Haematol.* 86: 804-809.
- Lee MR, Park SH, Choi YM, Jun JK, et al. (2009). Carrier detection and prenatal diagnosis of hemophilia A in a Korean population by fluorescent per analysis of dinucleotide repeats in intron 13 and intron 22 of the factor VIII gene. *Korean J. Obstet. Gynecol.* 52: 552-558.
- Levinson B, Kenwick S, Lakich D, Hammonds G Jr, et al. (1990). A transcribed gene in an intron of the human factor VIII gene. *Genomics* 7: 1-11.
- Levinson B, Kenwick S, Gamel P, Fisher K, et al. (1992). Evidence for a third transcript from the human factor VIII gene. *Genomics* 14: 585-589.
- Liu Q, Nozari G and Sommer SS (1998). Single-tube polymerase chain reaction for rapid diagnosis of the inversion hotspot of mutation in hemophilia A. *Blood* 92: 1458-1459.
- Naylor J, Brinke A, Hassock S, Green PM, et al. (1993). Characteristic mRNA abnormality found in half the patients with severe haemophilia A is due to large DNA inversions. *Hum. Mol. Genet.* 2: 1773-1778.
- Naylor JA, Buck D, Green P, Williamson H, et al. (1995). Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Hum. Mol. Genet.* 4: 1217-1224.
- Peake IR, Lillcrap DP, Boulyjenkov V, Briet E, et al. (1993). Haemophilia: strategies for carrier detection and prenatal diagnosis. *Bull. World Health Organ.* 71: 429-458.

- Richards B, Heilig R, Oberle I, Storzjohann L, et al. (1991). Rapid PCR analysis of the St14 (DXS52) VNTR. *Nucleic Acids Res.* 19: 1944.
- Rubinsztein DC, Amos W, Leggo J, Goodburn S, et al. (1995). Microsatellite evolution - evidence for directionality and variation in rate between species. *Nat. Genet.* 10: 337-343.
- Tuddenham EG, Schwaab R, Seehafer J, Millar DS, et al. (1994). Haemophilia A: database of nucleotide substitutions, deletions, insertions and rearrangements of the factor VIII gene, second edition. *Nucleic Acids Res.* 22: 4851-4868.
- Van de Water NS, May SJ and Browett PJ (1994). Possible misdiagnosis using the Xba I polymorphism for prenatal diagnosis of haemophilia A. *Br. J. Haematol.* 88: 613-614.
- Wion KL, Tuddenham EG and Lawn RM (1986). A new polymorphism in the factor VIII gene for prenatal diagnosis of hemophilia A. *Nucleic Acids Res.* 14: 4535-4542.