

# Remarkable variation in the informativeness of RFLP markers linked to hemophilia B locus in Indian population groups: Implication in the strategy for carrier detection

S. Mukherjee<sup>a</sup>, A. Saha<sup>a</sup>, Senthil Kumar P.<sup>a</sup>, G.R. Chandak<sup>b</sup>, P.P. Majumder<sup>c</sup> and K. Ray<sup>a,\*</sup>

<sup>a</sup>Human Genetics and Genomics and Division, Indian Institute of Chemical Biology, Kolkata, India

<sup>b</sup>Genome Research Group, Centre for Cellular and Molecular Biology, Hyderabad, India

<sup>c</sup>Indian Statistical Institute, Kolkata, India

**Abstract.** Hemophilia B, an X-linked recessive bleeding disorder, is caused by heterogeneous mutations in the factor IX (F9) gene. Hence, carriers of the disease are usually detected by F9 gene linked RFLP analysis. We aimed to test a set of RFLP markers (DdeI, XmnI, MnlI, TaqI & HhaI), used worldwide for carrier detection, to estimate its heterozygosity in different population groups of India, and identify additional single nucleotide polymorphisms (SNPs) if necessary. A total of 8 population groups encompassing different regions of India, consisting of 107 unrelated normal females without any history of hemophilia B in the family and 13 unrelated obligate carriers were recruited in the study. Regions of F9 gene were amplified by PCR from genomic DNA of the donors followed by restriction enzyme digestion and/or sequencing as appropriate. Combined informativeness for the markers varied between 52–86% among normal females belonging to different geographical locations of India. Haplotype analysis revealed that the most prevalent haplotype lacked the restriction sites for all five RFLP markers. Screening regions of F9 gene that harbor 10 SNPs reported in dbSNP yielded only two SNPs, which increased the overall informativeness in each population group and heterozygosity in the obligate carriers for the disease from 38% to 69%. Our data show that heterozygosity of commonly used RFLP markers is remarkably variable across different regions of India. Thus prudent selection of the markers based on specific population groups including usage of additional markers is recommended for efficient carrier detection.

Keywords: Carrier detection, F9, hemophilia B, marker, RFLP

## 1. Introduction

Hemophilia B is an X-linked recessive bleeding disorder affecting approximately 1 in 30,000 males and caused by deficient clotting factor IX (FIX) ac-

tivity [1]. FIX is a plasma thromboplastic component and plays an important role in blood coagulation cascade by activating factor X through interactions with calcium, membrane phospholipids and activated factor VIIIa by limited proteolysis [2]. Factor IX (F9) gene, located at Xq27, spans 33 kb of genomic DNA comprising 8 exons and encodes a 1.4 kb mRNA [3]. The FIX coagulant activity and immunoassay results have considerable overlap between the levels in carriers and normal women to unambiguously establish the disease status. Several studies have re-

---

\*Corresponding author: Dr. Kunal Ray, Human Genetics & Genomics Division, Indian Institute of Chemical Biology, 4 Raja S. C. Mullick Road, Jadavpur, Kolkata – 700 032, India. Tel.: +91 33 2473 3491 (or 0492 or 6793); +91 33 2473 5197 (or 2472-3967); E-mail: kunalray@gmail.com, thisiskr@rediffmail.com.

vealed high mutational heterogeneity and a relatively high rate of *de novo* mutations in F9 gene [4,5]. Over 500 mutations are recorded in the Hemophilia B mutation database (<http://www.kcl.ac.uk/ip/petergreen>), and strategies for identification of common mutations based on the database analysis have also been reported [6]. Mutational heterogeneity has also been reported in Indian patients [7,8]. It is estimated that number of hemophilia B patients in India is about 10,000 [9]. The mutational heterogeneity and the inability of the immunoassay results to detect carriers necessitated an indirect approach for DNA based carrier analysis and prenatal diagnosis in the affected families, which involves tracking the mutant chromosome by analysis of intragenic markers at hemophilia B locus. However, this strategy fails if the markers are not informative in the family members. In 1993, WHO published a list of markers [10] for efficient carrier detection based on early reports on high heterozygosity of these markers in Caucasian population. Although, limited studies have reported these markers to be highly heterozygous in western and north Indian population groups [9,11], our initial studies on a mixed population from eastern India indicated lower heterozygosity of these markers. Therefore, we undertook this study to examine the heterozygosity of these markers in different population groups across the country. We also report identification of two new F9 polymorphisms with high levels of heterozygosity, which improve the efficiency of carrier detection for hemophilia B.

## 2. Materials and methods

### 2.1. Study population

A total of 107 unrelated normal females from eight population groups were recruited in the study, which included three ethnic groups (Brahmin, Mahishya and Bagdi) and a mixed urban population from Kolkata, West Bengal from eastern region, three ethnic groups (Iyenger, Ulladan and Rajakoya) from Andhra Pradesh in southern region, and one ethnic group (Garasias) and a mixed urban population from Ahmedabad, Gujarat from western region of the country (Table 1). Institutional ethics committee of respective institutes approved the study following the guidelines of Indian Council of Medical Research, Govt. of India for handling of human samples. Our study population also contained 13 unrelated obligate carriers of Hemophilia B recruited through Hemophilia Federation of India,

comprising eight from Eastern India (Kolkata chapter) and five from Southern India (Chennai chapter). Blood samples were collected from all these individuals after obtaining their informed consent.

### 2.2. DNA analysis

Genomic DNA was isolated from each individual using the conventional phenol-chloroform method [12]. PCR was carried out in total volume of 25.0  $\mu$ l containing 50–100 ng genomic DNA, 0.4  $\mu$ M of each primer, 0.2 mM of each dNTP, 2.5 mM of MgCl<sub>2</sub> and 0.5 unit of Taq polymerase (Invitrogen, Carlsbad, CA). The RFLP markers (DdeI, XmnI, MnlI, TaqI & HhaI) were amplified using primers as described in WHO bulletin (1994) with a correction in the reverse primer for HhaI marker (5'-AGATTTCAAGCTACCAACAT-3' in place of 5'-AAGTACCTGCCAAGGGAATTGACCTGG-3'). The following conditions were used to amplify the region of interest (markers): initial denaturation at 94 °C for 3 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 52–58 °C for 30 s (based on the T<sub>m</sub> of the primer pairs) and extension at 72 °C for 30 s, and a final extension at 72 °C for 3 min. Aliquots of 5  $\mu$ l of the PCR product were checked for amplification and subjected to digestion using appropriate restriction enzymes under the conditions described by the manufacturer (New England Biolabs, Beverly, MA). The DNA fragments in the digest were separated by electrophoresis in polyacrylamide gel (6%) and visualized under UV light after ethidium bromide staining.

To identify new SNP, segments of introns 3, 4 and 5 of F9 gene were amplified. The PCR products were column purified using Quiagen PCR purification kits (Quiagen, Hilden, Germany) and bi-directional sequencing was performed using an ABI prism 3100 DNA sequencer with dye terminator chemistry. Nucleotide changes were detected by identifying double peaks in the chromatogram due to heterozygosity of the DNA sample. The sequences were analyzed using pairwise BLAST [13] to identify any changes from the normal sequences available in the database. Alteration of restriction site due to identified nucleotide variant was examined by WEBCUTTER (<http://www.firstmarket.com/cutter/cut2.html>). A SNP (A>C, rs422187) located in intron 5 was found to generate a site for Hpy188I and was analyzed by PCR-RFLP using primer pair: 5'-GGTCCTGGTGAATATGGCTGTG-3', and 5'-TACGGAAATAGAATAGGTGTTC-3'. However, another SNP (C>T, rs392959) in intron 3, did not alter any restriction site and was analyzed by di-

Table 1  
Population groups studied for heterozygosity of RFLP markers linked to hemophilia B locus

Region of India	Population group	Number of females
East	Brahmin, Bagdi, Mahishya ethnic groups and mixed population from Kolkata	69
West	Garasias, mixed population from Ahmedabad	21
South	Rajakoya, Iyenger and Ulladan	17

Study reported in Chowdhury et al. (2001) on ethnic groups of north Indian population has been used for comparison with other regions of India.

Table 2  
Allele frequency and heterozygosity of RFLP markers in Indian population groups

RFLP	EAST (69)				SOUTH (17)				WEST (21)				NORTH (28)*
	+	-	Heterozygosity		+	-	Heterozygosity		+	-	Heterozygosity		Heterozygosity
			Exp.	Obs.			Exp.	Obs.			Exp.	Obs.	
DdeI	0.086	0.928	0.134	0.145	0.118	0.882	0.208	0.235	0.310	0.690	0.427	0.476	0.50
TaqI	0.200	0.8	0.296	0.260	0.118	0.882	0.208	0.235	0.261	0.739	0.385	0.238	NA
Xmn I	0.071	0.929	0.194	0.188	0	1.000	–	0	0.238	0.762	0.367	0.190	0.23
Mnl I	0.114	0.886	0.238	0.217	0.118	0.882	0.208	0.235	0.214	0.786	0.336	0.330	NA
Hha I	0.100	0.9	0.182	0.203	0.235	0.765	0.3595	0.470	0.190	0.810	0.308	0.380	0.46
Hpy188I	0.157	0.842	0.267	0.232	0.176	0.835	0.294	0.352	0.175	0.825	0.288	0.333	NA

\*Refers to a separate study from North India (Chowdhury et al. 2001); NA, Not available.

Note: Figures in the parentheses refer to the number of female samples taken from each population. In case of Southern India, with lowest number of female donors, we also examined 73 males for the same set of markers and observed similar frequencies except for HhaI (0.123 vs 0.235).

rect sequencing of the amplicon generated using the primer pair 5'-GTCTCTTGTTGTATTTGACCCCA-3' and 5'-GTCAGAATGAGAAGGGAATC-3'.

### 2.3. Statistical analysis

The allele frequency for each marker was calculated using the allele counting method, which was then used to estimate the expected heterozygosity. Haplotype frequencies and linkage disequilibrium between different markers were estimated following Expectation-maximization algorithm using ARLEQUIN software (version 2.1).

## 3. Results

To evaluate the utility of the commonly used RFLP markers (DdeI, XmnI, TaqI, MnlI, HhaI) in F9 gene for carrier detection, we first determined heterozygosity of each marker in the selected population groups. Then, relative and overall informativeness of these markers was determined. Since the overall heterozygosity for all five markers did not reach 100%, we also attempted to identify markers with high heterozygosity in the F9 gene. Haplotype analysis was performed to find variation in chromosomes in each population group. Two additional SNPs were identified which increased the efficiency of hemophilia B carrier detection as demonstrated by genotyping the obligate carriers.

### 3.1. Highly variable heterozygosity of RFLP markers among population groups

The heterozygosity of the RFLP markers showed remarkable heterogeneity among different population groups of India (Table 2). While the XmnI RFLP was completely monomorphic, HhaI proved to be the most efficient marker in the southern group with a heterozygosity of 0.47. One of the ethnic groups of southern India (Ulladan), was observed to be completely monomorphic for three RFLPs (TaqI, XmnI & MnlI) (data not shown separately). Analysis of 73 males from the Southern India revealed frequencies for the markers to be similar to the females from the same population group despite relatively small number of individuals available for analysis. Similarly, the eastern group also showed low heterozygosity for most of the markers, with TaqI being the most efficient marker with highest heterozygosity (0.26) followed by MnlI (0.22) and HhaI (0.20). Compared to the eastern and southern groups, the western group showed relatively higher heterozygosity for all the markers, which is consistent with the previous reports [9,11]. In this group, DdeI was the most heterozygous marker (0.48) followed by HhaI (0.38) and MnlI (0.33). Heterozygosity for these markers has been reported to be quite high in the north Indian population (Chowdhury et al., 2001). Since we did not have any female sample from northern In-

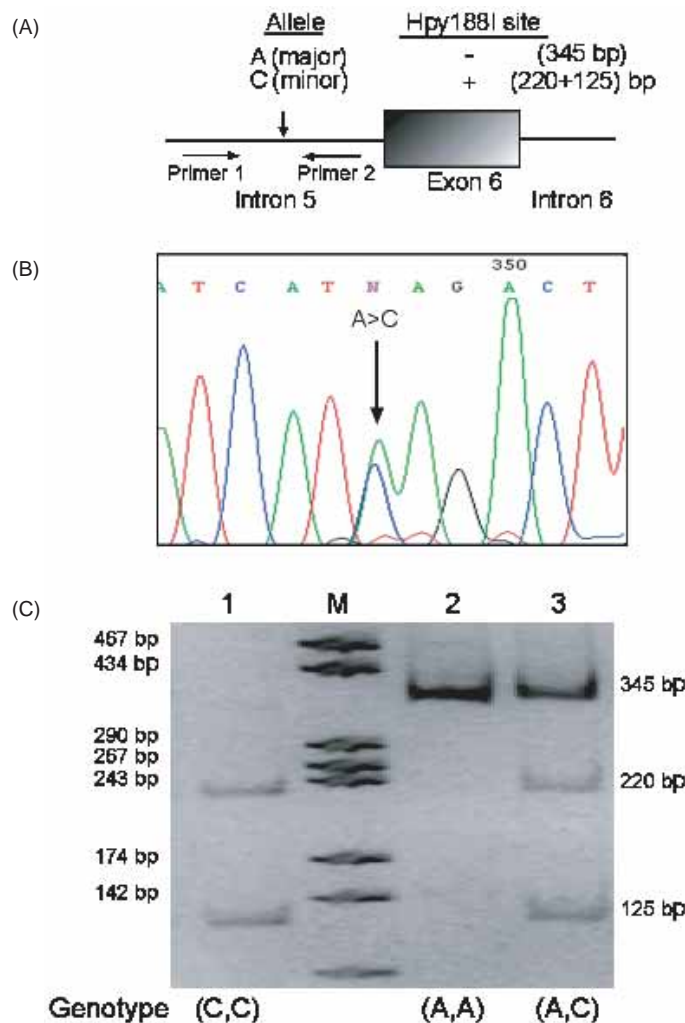


Fig. 1. Detection of SNP in intron 5. (A) Location of the SNP (A>C, rs422187) within F9 gene and the nature of the nucleotide change are shown. The allelic difference resulting from the SNP was determined by Hpy188I digestion of the PCR product obtained using a pair of primers selected from intron 5 of the gene. (B) Representative electropherogram showing the location of the SNP as a double peak in the heterozygous condition (arrow). (C) 6% PAGE of PCR products from three representative samples and their genotype are shown. Lane M is pBS/HaeIII digested DNA molecular weight marker. The size of the DNA molecular weight marker and the Hpy188I digested DNA fragments are shown on the sides of the gel.

dia, we determined the allele frequencies of these five markers in 30 male samples and also found them to be relatively high (DdeI = 0.33, XmnI = 0.10, HhaI = 0.172, 30 chromosomes) and comparable (DdeI = 0.40, XmnI = 0.18, HhaI = 0.36, 17 chromosomes) to the reported study (Chowdhury et al., 2001). Three of the commonly used markers (TaqI, XmnI and MnlI) were observed to be in linkage disequilibrium in all three population groups ( $p \leq 0.002$ ) except for northern population where LD was observed between TaqI and XmnI only ( $p = 0.001$ ). The combined heterozygosity for these markers was 52%, 59%, 86% for eastern,

southern and western Indian populations, respectively (Fig. 2).

### 3.2. Estimation of heterozygosity for common RFLP markers among obligate carriers for hemophilia B

In addition to determining the heterozygosity of the common RFLP markers in different population groups in India, we calculated the heterozygosity of these markers in 13 unrelated obligate carriers of hemophilia B. Only 5 carriers (38%) were found to be heterozy-

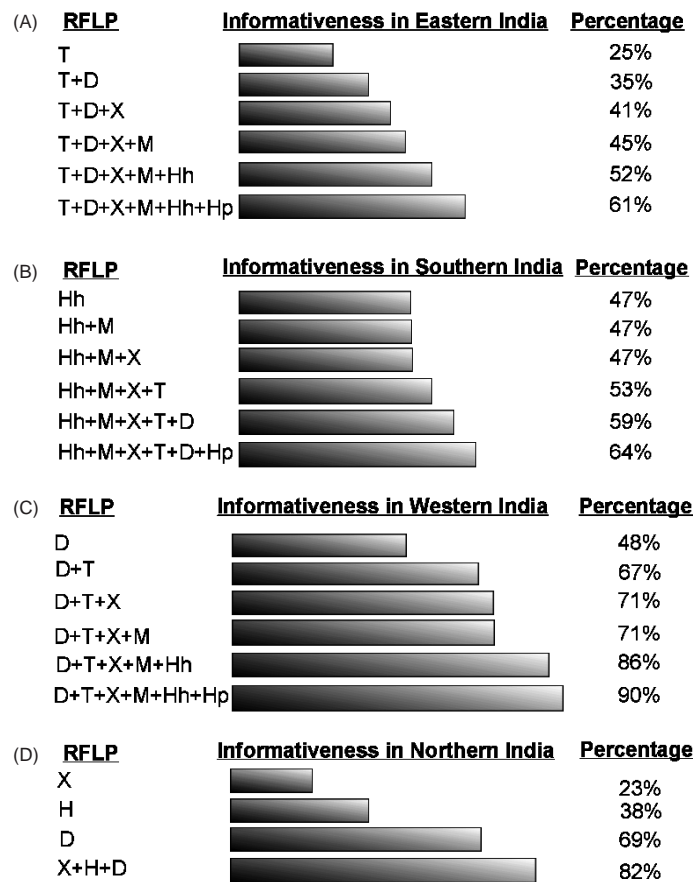


Fig. 2. Informativeness of the RFLP markers in different population groups of India. Panels A, B and C respectively represent the Eastern, Southern and Western regions. RFLP markers analyzed in each bar diagram is denoted by DdeI (D), HhaI (Hh), Hpy188I (Hp), MnlI (M), TaqI (T), and XmnI (X). Panel D shows informativeness for North India obtained from a separate study (Chowdhury et al. 2001). Informativeness for each marker was calculated as the percentage of females heterozygous in each cluster.

gous using these five common RFLP markers. Despite the fact that only a modest number of obligate carriers were analyzed, the observation made based on studies on normal individuals provides a wholesome picture on the genomic variability in the hemophilia B locus.

### 3.3. Identification of new SNP markers and assessment of their heterozygosity in various population groups

We selected multiple regions of F9 gene harboring at least 10 SNPs as recorded in dbSNP ([www.ncbi.nlm.nih.gov/SNP](http://www.ncbi.nlm.nih.gov/SNP)). Sequence analysis of these regions from 35 normal unrelated females revealed only two SNPs namely C>T in intron 3 (rs392959) and A>C in intron 5 (rs422187) of F9. The first SNP did not alter any restriction site and was scored in obligate carriers by direct sequencing. Two obligate carriers,

homozygous for other markers, were heterozygous for this SNP. For the lack of an easy format assay to score the alleles, we did not estimate heterozygosity of this marker in the normal population consisting of different ethnic groups.

The second SNP (i.e. A>C in intron 5), however, generated an Hpy188I restriction enzyme site. A 345 bp amplicon encompassing the Hpy188I site was generated from genomic DNA, which would cleave into two DNA fragments (220 bp and 125. bp) in the presence of the variant allele (Fig. 1). The heterozygosity of this marker was observed to be fairly high across all population groups (Table 2). The allele frequency of the marker was also observed to be high in the group of 30 males from northern India (0.416). Using this marker, two additional obligate carriers for hemophilia B were found to be heterozygous, thus increasing the overall efficiency of carrier detection to 69% from the existing 38% using the common RFLP markers.

Table 3  
Major haplotype frequencies in Indian population groups

S. No.	Haplotypes	East (69)	South (17)	West (21)
A	1 1 1 1 1 1	61.4%	69.6%	30.8%
B	1 1 1 1 2 1	4.2%	7.0%	8.4%
C	1 2 1 1 1 1	7.5%	0%	0%
D	1 2 2 2 1 1	0%	0%	15.0%
E	2 1 1 1 1 1	1.6%	3.7%	15.9%
F	2 1 1 1 2 1	1.5%	3.0%	7.4%
G	1 1 1 1 1 2	6.5%	3.3%	7.5%

Number in the parentheses refers to the number of female samples taken from each population. Only those markers have been included which have more than 5% frequency in any of the population groups. Order of the markers: DdeI-TaqI-XmnI-MnlI-HhaI-Hpy188I. For each RFLP locus allele 1 and allele 2 represent absence and presence of the polymorphic restriction site, respectively.

### 3.4. Informativeness of the markers in the population groups

To estimate the relative efficiency of each RFLP marker and their collective efficiency for carrier detection in different population groups, informativeness of the markers was determined based on the heterozygosity values. It was observed that in different populations, the priority of choice of marker would vary widely (Fig. 2). For example TaqI would be the marker of choice for eastern India, but HhaI and DdeI would be the markers to be used first for southern and western Indian groups respectively. We estimated that the new marker (Hpy188I) would increase the informativeness marginally from 52% to 61%, 59% to 64% and 86% to 90% in eastern, southern and western groups, respectively (Fig. 2). Two newly identified SNPs together were observed to increase the carrier detection efficiency from 38% to 69% (5/13 to 9/13) among the pool of obligate carriers. However, the increase in the informativeness using the new markers in each population group was only marginal.

### 3.5. Haplotype analysis

We estimated the number of major haplotypes and the frequencies in each population group using the RFLP markers. Thus, the genotypes at all six RFLP markers were used to construct haplotypes of female samples in each cluster (Table 3). Seven major haplotypes, representing > 5% of chromosomes, were identified. The frequency of haplotype A, lacking restriction site for all six RFLP varied from 60–70% among the eastern and southern population groups, while the western group had 31% of the same haplotype (Table 3). Due to lack of female samples, north Indian male samples were examined to identify the common

haplotypes and found to harbor 44.4% of the haplotype A. This observation suggested that the eastern and southern population groups are likely to be less informative using the RFLP markers compared to the population groups from the northern and western part of India.

## 4. Discussion

The prevalence of hemophilia B is similar worldwide and the disease is caused by frequent *de novo* mutations spanning throughout the F9 gene, which is 30–50% of all patients [4,5].

For all practical purposes, the load of mutation in the population cannot be eliminated but needs to be controlled by mutation screening and carrier detection. Although carrier analysis for relatively smaller gene could be done by direct sequencing and is practiced more with lowering the cost of large scale sequencing, the strategy is not ideal for a wider use for diagnostic purposes in developing countries with less resources available for such endeavor. Hence carrier detection by linkage analysis using informative markers is the best alternative choice.

The PCR based RFLP markers commonly advocated for use in carrier detection of hemophilia B were originally based on the studies primarily conducted on the Caucasians. Subsequently, these markers have been reported to have low heterozygosity in the Chinese [10, 14,15], Japanese [16,17] and Malay population [10]. Indians represent a large proportion of the world population and the population genetic studies suggest multiple ethnic groups, which remain relatively isolated through marriage within the community. Therefore, it is likely that the polymorphic nucleotides in F9 gene would have considerable variation in the entire country. In this context, our finding that the markers linked to hemophilia B locus have variable informativeness in different population groups is not unexpected. This precise information, however, is important for informed decision regarding marker selection for carrier detection in specific population group. Haplotype analysis shows that a large proportion of the chromosome is completely monomorphic for the five RFLP markers commonly chosen for carrier analysis and underscores importance of identification of new markers. Our attempt to conduct the study on a larger sample size was particularly challenging since only female samples were required to test heterozygosity of F9. Due to socio-economic and cultural reasons, in many eth-

nic groups, males are usually more easily available and persuaded as the donors for blood collection in such studies. Therefore, while studies involving a much larger number of normal and carrier females would certainly be preferable, the present study clearly shows remarkable difference in the informativeness of the markers studied and the results are consistent with previous studies undertaken in some ethnic groups as mentioned before.

Identification of new markers needs immense efforts as our initial investigation identified only two out of 10 reported SNPs to be of use in carrier analysis. We did not detect any novel SNP, but identified two SNPs significantly increased the efficiency of the carrier detection. This observation is also consistent with the fact that occurrence of polymorphism in X chromosome is relatively rare [18]. Our observations highlight the need to develop new markers with high heterozygosity in Indian population for successful carrier detection and for reducing the burden of hemophilia B among Indians. The same may also hold for other populations since the commonly used markers are reported to have low heterozygosity in many of them [11,15,18]. However, in the best case scenario there will be a certain proportion of obligate carriers left, who will still be uninformative even with the use of these markers due to the LD between different markers of F9 gene and direct mutation detection will be the only way to perform carrier detection in such cases.

The information content of the SNP markers in a specific region of the genome is dependent on determined haplotype blocks based on linkage disequilibrium between different markers. This has been best demonstrated recently through the HAPMAP project (<http://www.hapmap.org/> [19]). However the project does not include population groups from Indian subcontinent, which makes our study relevant despite relatively small size of samples as appropriate for this part of the work. However the studies are underway to decipher the landscape of genomic variation in Indian population [20], which will shed more light on the utility of the commonly used markers for detection of hemophilia B carriers.

### Acknowledgements

The authors thank the patients who participated in the study and Dr. Amal Saha and Mr. Ujjal Roy of the Haemophilia Federation of India. The study has been supported by the Department of Science and Technol-

ogy (DST), Govt. of India and the Council of Scientific and Industrial Research (CSIR), Govt. of India (project no. CMM-0016). SM is a recipient of a Research Fellowship from CSIR.

### References

- [1] F. Giannelli, P.M. Green, S.S. Sommer, D.P. Lillicrap, M. Ludwig, R. Schwaab, P.H. Reitsma, M. Goossens, A. Yoshioka and G.G. Brownlee, Haemophilia B: database of point mutations and short additions and deletions, fifth edition, 1994, *Nucleic Acids Res* **22** (1994), 3534–3546.
- [2] K.G. Mann, M.E. Nesheim, W.R. Church, P. Haley and S. Krishnaswamy, Surface-dependent reactions of the vitamin K-dependent enzyme complexes, *Blood* **76** (1990), 1–16.
- [3] S. Yoshitake, B.G. Schach, D.C. Foster, E.W. Davie and K. Kurachi, Nucleotide sequence of the gene for human factor IX (antihemophilic factor B), *Biochemistry* **24** (1985), 3736–3750.
- [4] D.D. Koeberl, C.D. Bottema, R.P. Ketterling, P.J. Bridge, D.P. Lillicrap and S.S. Sommer, Mutations causing hemophilia B: direct estimate of the underlying rates of spontaneous germline transitions, transversions, and deletions in a human gene, *Am J Hum Genet* **47** (1990), 202–217.
- [5] D.D. Koeberl, C.D. Bottema, J.M. Buerstedde and S.S. Sommer, Functionally important regions of the factor IX gene have a low rate of polymorphism and a high rate of mutation in the dinucleotide CpG, *Am J Hum Genet* **45** (1989), 448–457.
- [6] S. Mukherjee, A. Mukhopadhyay, K. Chaudhuri and K. Ray, Analysis of haemophilia B database and strategies for identification of common point mutations in the factor IX gene, *Haemophilia* **9** (2003), 187–192.
- [7] S. Mukherjee, A. Mukhopadhyay, D. Banerjee, G.R. Chandak and K. Ray, Molecular pathology of haemophilia B: identification of five novel mutations including a LINE 1 insertion in Indian patients, *Haemophilia* **10** (2004), 259–263.
- [8] G. Jayandharan, R.V. Shaji, M. Chandy and A. Srivastava, Identification of factor IX gene defects using a multiplex PCR and CSGE strategy—a first report, *J Thromb Haemost* **1** (2003), 2051–2054.
- [9] S. Shetty, K. Ghosh, A. Pathare, R. Colah, S. Badakare and D. Mohanty, Factor VIII and IX gene polymorphisms and carrier analysis in Indian population, *Am J Hematol* **54** (1997), 271–275.
- [10] I.R. Peake, D.P. Lillicrap, V. Boulyjenkov, E. Briet, V. Chan, E.K. Ginter, E.M. Kraus, R. Ljung, P.M. Mannucci, K. Nicolaidis et al., Haemophilia: strategies for carrier detection and prenatal diagnosis, *Bull World Health Organ* **71** (1993), 429–458.
- [11] M.R. Chowdhury, M. Kabra and P.S. Menon, Factor IX gene polymorphisms in Indian population, *Am J Hematol* **68** (2001), 246–248.
- [12] J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning, A Laboratory Manual*, 1989.
- [13] T.A. Tatusova and T.L. Madden, BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences, *FEMS Microbiol Lett* **174** (1999), 247–250.
- [14] C. de la Salle, Q. Wu, M.J. Baas, A. Hanauer, C. Ruan and J.P. Cazenave, Common intragenic and extragenic polymorphisms of blood coagulation factors VIII and IX are different in Chinese and Caucasian populations, *Clin Genet* **38** (1990), 434–440.

- [15] J.B. Graham, G.R. Kunkel, N.K. Egilmez, A. Wallmark, D.M. Fowlkes and S.T. Lord, The varying frequencies of five DNA polymorphisms of X-linked coagulant factor IX in eight ethnic groups, *Am J Hum Genet* **49** (1991), 537–544.
- [16] T. Kojima, M. Tanimoto, T. Kamiya, Y. Obata, T. Takahashi, R. Ohno, K. Kurachi and H. Saito, Possible absence of common polymorphisms in coagulation factor IX gene in Japanese subjects, *Blood* **69** (1987), 349–352.
- [17] H. Toyozumi, T. Kojima, T. Matsushita, M. Hamaguchi, M. Tanimoto and H. Saito, Diagnosis of hemophilia B carriers using two novel dinucleotide polymorphisms and Hha I RFLP of the factor IX gene in Japanese subjects, *Thromb Haemost* **74** (1995), 1009–1014.
- [18] H. Kaessmann, F. Heissig, A. von Haeseler and S. Paabo, DNA sequence variation in a non-coding region of low recombination on the human X chromosome, *Nat Genet* **22** (1999), 78–81.
- [19] The International Hapmap Consortium. A haplotype map of human genome, *Nature* **27** (2005), 1299–1321.
- [20] The Indian Genome Variation Consortium. The Indian genome variation database (IGVdb): a project overview, *Hum genet* **118** (2005), 1–11.