British Journal of Haematology, 1998, **100**, 589–593

Analysis of the two microsatellite repeat polymorphisms of the factor VIII gene in the Turkish population

HAMDİ JARJANAZİ, A. ANIL TIMUR, OSMAN EL-MAARRİ AND S. HANDE ÇAĞLAYAN
Department of Molecular Biology and Genetics, Boğaziçi University, Bebek, İstanbul, Turkey

Received 17 August 1997; accepted for publication 25 November 1997

Summary. DNA-based diagnosis of haemophilia A has previously been carried out by linkage analysis using two highly informative markers, Hind III RFLP and St14 VNTR, for affected Turkish families. In the present study the number and frequency of the microsatellite alleles at introns 13 and 22 in the factor VIII (FVIII) gene were analysed in order to increase the rate of informative females and accuracy of linkage analysis. Six alleles were observed at both loci. The two most frequent alleles of each locus were the same as the two common alleles found in Anglo-Americans. The comparison of heterozygosity of both microsatellite loci showed that the Turkish population is slightly less polymorphic than Anglo-Americans but more polymorphic than Chinese, Slavs and Uzbekians. The additional use of the two microsatellite repeat polymorphisms with the previously established informative markers has been accepted as the most effective strategy in DNA diagnosis by linkage analysis for the assessment of haemophilia A carriers and affected fetuses in the Turkish population. The modifications adopted in this study for the multiplex PCR analysis of the microsatellite repeat polymorphism eliminated the use of radioactivity and sequencing gels, reducing cost and labour.

Keywords: haemophilia A, microsatellite repeat alleles, multiplex PCR, genetic counselling, Turkish population.

Haemophilia A is an X-linked bleeding disorder which is the most frequently encountered, with an incidence of 1/5000 (Hoyer, 1994). The disease results from mutations in the factor VIII (FVIII) gene which is located in the Xq28 region spanning 186 kb and consisting of 26 exons that code for 9 kb mRNA (Gitschier *et al*, 1984). Genetic counselling for haemophilia A has been widely carried out for more than a decade by the use of DNA linkage markers (Broker-Vriends *et al*, 1992) and several intragenic RFLPs have been identified and used in linkage analysis (Green *et al*, 1991). The presence of tight linkage between some markers and the unavailability of PCR primers for some restriction sites, limits the use of RFLPs in tracking the defective allele (Peake *et al*, 1993). It has been observed that the Turkish population is as polymorphic as Anglo-Americans at the HindIII RFLP site or BclI RFLP site but the heterozygosity at the AlwNI site in intron 7 is significantly lower than Caucasians (Çağlayan *et al*, 1995b). In a comprehensive study of haemophilia A families in Turkey it has been shown that, by the combined use of HindIII, AlwNI and the extragenic VNTR marker St14, 86% of females were eligible for DNA linkage analysis (Çağlayan *et al*, 1995a).

The use of additional informative intragenic markers in DNA linkage analysis would certainly increase the accuracy in diagnosis. The $(CA)_n$ dinucleotide repeat polymorphism characterized by Lalloz *et al* (1991, 1994) in introns 13 and intron 22 of the FVIII gene were good candidates as multiallelic markers to be used in DNA linkage analysis. The allele frequency of these markers has been determined in Caucasian, Canadian, Italian, Thai, Slav, Uzbekian and Chinese populations (Lalloz *et al*, 1991; Windsor *et al*, 1994; Goodeve *et al*, 1996; Aseev *et al*, 1994; Yip *et al*, 1994). A multiplex PCR analysis of the two intragenic microsatellite repeat polymorphisms has been introduced for convenience in the genetic diagnosis of haemophilia A. The combination of the multiplex PCR analysis of the two dinucleotide repeats with the PCR analysis of intron 18 Bcl I RFLP has been suggested as the most effective strategy for initiating haemophilia A genetic testing. This protocol provided informative results in approximately 88% of families (Windsor *et al*, 1994).

We aimed to establish whether the dinucleotide repeats at introns 13 and 22 are also polymorphic in the Turkish population and if their combined use with HindIII RFLP and

Correspondence: Dr S. Hande Çağlayan, Department of Molecular Biology and Genetics, Boğaziçi University, 80815 Bebek, İstanbul, Turkey.

St14 VNTR would increase the number of informative families for prenatal testing or carrier determinations in the Turkish population.

MATERIALS AND METHODS

Materials. *Thermus aquaticus* (Taq) DNA polymerase, AMV reverse transcriptase and T7RNA polymerase were from Promega; DIG detection kit and Nylon membrane were from Boehringer, γ-(^{32}P) from Amersham. Blood samples from Turkish haemophilia A patients and their relatives were provided by haematology clinics in all parts of the country.

Study population. 397 and 252 chromosomes from unrelated Turkish individuals, including some unrelated members of haemophilia A families, were used to determine the frequency of intron 13 and intron 22 microsatellite repeat alleles in the population. 10 ml of blood was drawn into tubes containing EDTA, and leucocyte DNA was isolated by the NaCl extraction method (Miller *et al*, 1988).

Determination of the length of CA repeats. The length of the CA repeats at intron 13 was determined using GAWTS (Genomic Amplification With Trancript Sequencing) as the sequencing method (Stoflet *et al*, 1988). The primer sequences used to amplify the microsatellite repeat region at intron 13 were as described in Lalloz *et al* (1991) except that the phage T7 promoter sequence was added to the 5′ end of the reverse primer. The sequence of the primers was as follows: forward primer: 5′-TGCATCACTGTACATATGTATCTT; reverse primer: 5′-GGTACCTAATACGACTCACTATAGGGAG-ACCAAATTACATATGAATAAGCC.

PCR amplification reactions were carried out in a total volume of 25 μl, containing 2·5 μl of 10×Promega Mg^{2+} free buffer, 6·25 pmol of each primer, 3·5 mM $MgCl_2$, 1 mM of each of the four deoxyribonucleoside triphosphates, 1 unit of Taq polymerase and 200–300 ng of genomic DNA. The samples were denatured for 5 min at 94°C and amplified in 35 cycles involving 35 s denaturation at 94°C, 1 min annealing at 52°C and 1 min synthesis at 72°C. After the last cycle the synthesis was extended for 5 min. The amplified region was transcribed using T7 RNA polymerase and sequenced using γ-(^{32}P) end-labelled forward primer and reverse transcriptase, to determine the number of CA repeats in the amplified region.

Determination of the frequency of microsatellite repeat alleles. The PCR products of the intron 13 $(CA)_n$ region from 397 unrelated individuals were denatured and run on a polyacrylamide sequencing gel. Two alleles (CA_{20} and CA_{22}) the size of which were determined by direct sequencing described previously were chosen as size markers. The primer sequences used for PCR were: forward primer: 5′-TGCAT-CACTGTACATATGTATCTT; reverse primer: 5′-CCAAAT-TACATATGAATAAGCC (Lalloz *et al*, 1991). The piece of the gel that contained the PCR products was blotted onto positively charged nylon membrane by capillary blotting. The membrane was then hybridized to a DIG-11-dUTP intron 13 $(CA)_n$ repeat labelled probe for 3 h and then incubated in a solution containing anti-DIG alkaline phosphatase at a concentration of 150 mU/ml. The colour detection was performed by incubating the membrane in a solution containing 0·45 μl/ml nitroblue tetrazolium (NBT) solution and 0·35 μl/ml X-phosphate in the dark until the formation of colour precipitate was complete.

To determine the number and frequency of the intron 22 microsatellite repeat the PCR products from 252 unrelated individuals were run on 15×16 cm 7% polyacrylamide gels and stained with silver nitrate as described below. The primer sequences used to amplify intron 22 microsatellite repeats were: forward primer: 5′-TTCTAAGAATGTAGTGTGTGTGTG; reverse primer: 5′-TAATGCCCACATTATAGA (Lalloz *et al*, 1994).

Multiplex PCR and silver staining of the products. Multiplex PCR conditions were as described by Windsor *et al* (1994). The same PCR primers were used as given above. 6 μl of PCR products were mixed with 5 μl of loading buffer (bromophenol blue and xylene cyanol) and run on 7% polyacrylamide gels (16×15 cm) at 500 V until the xylene die reached the bottom of the gel.

The gel was stained by using an ultrasensitive staining method of nucleic acids with silver nitrate with some modifications (Beidler *et al*, 1982). The gel was fixed for 3 min in 10% ethanol and 0·5% acetic acid twice and stained with 0·1% $AgNO_3$ in daylight for 10 min, washed twice with distilled H_2O for 10 s, and developed in 1·5% NaOH and 0·01% $NaBH_4$ and 0·015% formaldehyde for 20 min. The reaction was stopped by 0·75% Na_2CO_3 for 7 min and then incubated in 5% glycerol for 5 min. After discarding the extra fluid the gel was sealed in a nylon bag for storage.

RESULTS

Sequence analysis of the intron 13 microsatellite repeat was performed on 13 individuals to determine the length of the CA repeats and four different microsatellite repeat alleles, $(CA)_{20}$–$(CA)_{23}$, were observed (data not shown). Two of the

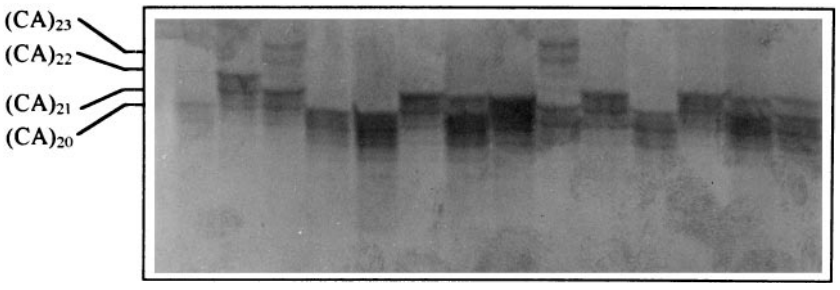

Fig 1. Four of the intron 13 $(CA)_n$ alleles, $(CA)_{20}$–$(CA)_{23}$, run on sequencing gels and detected by DIG labelling and colour detection system.

Table I. The number of intron 13 microsatellite repeat alleles and their frequencies in the Turkish population compared to the frequency of microsatellite repeat alleles in Caucasian, Canadian, Chinese, Slav and Uzbekian populations.

No. of $(CA)_n$ at intron 13	Allele frequency (%)							
	Caucasian[1]	Canadian[2]	Chinese[3]	Slav[4]	Uzbeki[4]	Italian[5]	Thai[5]	Turkish[6]
16	0·5	–	–	–	–	–	–	–
17	–	1·7	–	–	–	–	–	–
18	0·5	1·2	–	–	–	–	–	–
19	7	4	–	3	–	5·6	20·5	1·8
20	45	7	–	56	80	61	59·1	51
21	29	25	–	19	14	22·2	11·4	34
22	11	48	1	11	3	2·8	9·0	7·8
23	5	13	1	10	3	5·6	–	4·3
24	1	0·4	9	–	–	2·8	–	0·8
25	–	–	53	–	–	–	–	–
26	–	–	32	–	–	–	–	–
27	–	–	3	–	–	–	–	–
28	–	–	1	–	–	–	–	–
Expected heterozygosity (%)*	69	68	61	63	34	57·1	58·8	62
Observed heterozygosity (%)	91	71	53·7	34	–†	–†	–†	50
χ^2	7·01	0·13	0·87	13·35				2·32
P	0·001–0·01	0·07–0·9	0·5–0·3	<0·001				0·1–0·2

[1] Lalloz *et al*, 1991; [2] Windsor *et al*, 1994; [3] Yip *et al*, 1994; [4] Aseev *et al*, 1994; [5] Goodeve *et al*, 1996; [6] data obtained from the analysis of 397 independent X chromosomes in the present study.

*Expected heterozygosity $= 1 - (p^2 + q^2 + r \ldots)$, where p, q and r, etc., are allele frequencies.

† Observed heterozygosity rates for Uzbeki, Italians and Thai were not given in the original articles.

samples with a repeat length of $(CA)_{20}$ and $(CA)_{22}$ were selected as size markers to determine the frequency and number of alleles in the population. The markers were run with unknown samples on sequencing gels and probed with DIG-11 dUTP-labelled CA sequence. In this way, intron 13 microsatellite repeat alleles of 397 independent chromosomes were examined and six different alleles, $(CA)_{19}$–$(CA)_{24}$, were observed. Four of the intron 13 CA alleles are shown in Fig 1. The frequency of each allele is shown in Table I. $(CA)_{20}$ and $(CA)_{21}$ were the most frequent alleles with 51% and 34% frequency, respectively. $(CA)_{22}$ and $(CA)_{23}$ were less frequent alleles, whereas $(CA)_{19}$ and $(CA)_{24}$ were the two rarest alleles.

Intron 22 microsatellite repeat alleles were also examined on 252 independent X chromosomes and six alleles, $(CA)_{23}$–$(CA)_{28}$, were observed. The size of the alleles were determined by considering that the most frequent allele observed in the Turkish population to be equivalent to the most frequent allele observed in both Caucasian and Chinese populations. This allele has 26 (CA)(TC) repeats and has a frequency of 56% in the Turkish population (Table II). The second common allele was $(CA)_{25}$ with a frequency of 38·8%. $(CA)_{23}$, $(CA)_{24}$, $(CA)_{27}$, and $(CA)_{28}$ were also observed as rare alleles.

Expected heterozygosity was calculated, and χ^2 analysis of the observed and expected heterozygosity rates were carried out for both microsatellite repeat loci in the populations included for comparisons. The P values indicated that the observed heterozygosity for Canadians, Turks and Chinese

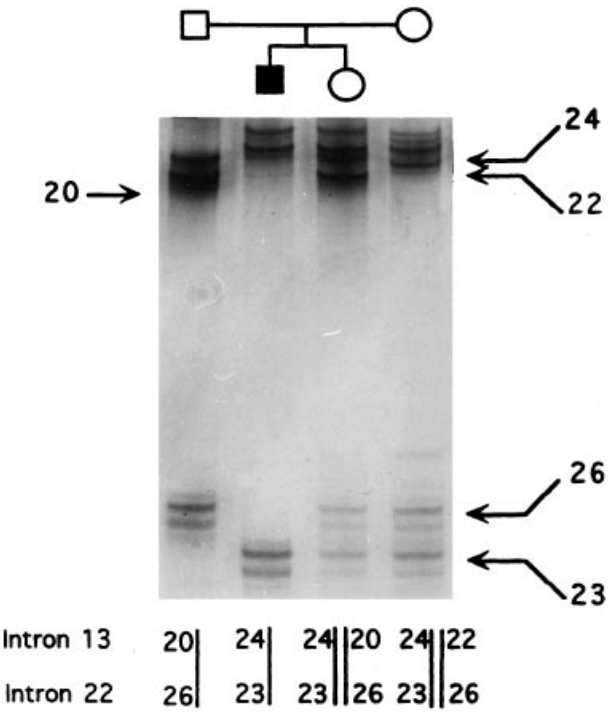

Fig 2. Multiplex analysis of the microsatellite repeats using silver staining.

Table II. Comparison of the frequency of microsatellite repeat alleles at intron 22 in Caucasian, Canadian and Chinese populations.

No. of $(CA)_n(TC)_n$ at intron 22	Allele frequency (%)					
	Caucasian[1]	Canadian[2]	Chinese[3]	Italian[4]	Thai[4]	Turkish[5]
20	–	–	1	–	–	–
21	–	–	–	–	–	–
22	–	–	–	–	–	–
23	1·3	6	–	2·5	–	0·4
24	–	53	–	5·0	–	1·2
25	30·7	35	29	22·5	30	38·8
26	66·7	3	63	65	66	56
27	–	2	7	5	4	3·2
28	1·3	–	–	–	–	0·4
Expected heterozygosity (%)*	46	59	51	52·1	47·3	53
Observed heterozygosity (%)	37	71	44·4	–†	–†	48·5
χ^2	1·76	2·44	0·85			0·38
P	0·1–0·2	0·1–0·2	0·3–0·5			0·5–0·7

[1] Lalloz *et al*, 1991; [2] Windsor *et al*, 1994; [3] Yip *et al*, 1994; [4] Goodeve *et al*, 1996; [5] data obtained from the analysis of 252 independent X chromosomes in the present study.

* Expected heterozygosity $= 1 - (p^2 + q^2 + r^2 \ldots)$, where p, q and r, etc., are allele frequencies.

† Observed heterozygosity rates for Italians and Thais were not given in the original articles.

were statistically significant for the intron 13 locus (Table I). At the intron 22 locus the expected heterozygosities were around 50% for four populations included for comparative study and the observed heterozygosity rates were statistically significant (Table II).

Multiplex analysis of the two microsatellite loci on sequencing gels are widely used in molecular diagnosis of haemophilia A and the genetic counselling involved. In an attempt to facilitate the use of multiplex analysis of the highly polymorphic microsatellite repeat polymorphisms, we used a smaller gel system and adopted the silver nitrate method to stain the DNA. Fig 2 shows the pattern of microsatellite repeat alleles in a multiplex PCR analysis followed by silver staining in members of a haemophilia A family.

A sample of 45 unrelated females from haemophilia A families were used to show the usefulness of the three intragenic loci in the molecular diagnosis of haemophilia A in the Turkish population. It was shown that using the three intragenic loci in linkage analysis, 64·5% of the females would be informative and therefore eligible for carrier identification and prenatal diagnosis. When the polymorphic St14 loci linked to the FVIII gene was also included in the linkage analysis the percentage of the informative females was increased to 93·3. The previous carrier diagnosis carried out for these families using HindIII RFLP and St14 VNTR was confirmed using $(CA)_n$ repeats.

DISCUSSION

Microsatellite loci have multiple alleles because new alleles can be formed more frequently in comparison to RFLP alleles due to mechanisms such as slippage and unequal crossing over. Multiallelic loci naturally have higher heterozygosity rates than biallelic loci and are better candidates as markers in linkage analysis. The use of all three intragenic loci spanning the FVIII gene (intron 13, intron 19, and intron 22 the largest intron) increased the heterozygosity rate in haemophilia A families, showing that the microsatellite loci will be useful markers in future linkage analysis studies.

The discrepancy in the length of the two common intron 13 and intron 22 alleles in the Canadian population may be explained by a founder effect. A founder would have carried intron 13 $(CA)_{22}$ and intron 22 $(CA)_{24}$ haplotype that increased in frequency in the Canadian population. Uzbeks, being an isolated population, have much less heterozygosity. The presence of the $(CA)_{20}$ in 80% of the Uzbeki population indicates that it is the original allele at inton 13 microsatellite loci.

In conclusion, the Turkish population is found to be genotypically closer to Caucasian populations than Asian, confirming the previous finding about the distribution of the FIX polymorphisms (Çağlayan *et al*, 1995b).

ACKNOWLEDGMENTS

We thank Dr Aytemiz Gürgey, Dr Gülten Aktuğlu and Dr Kaan Kavaklı for providing samples from the haemophilia A families. Hamdi Jarjanazi and Osman El-Maarri are TÜBİTAK (Turkish Scientific and Research Council) fellows. This study was supported by Boğaziçi University fund projects 94B0141, 95B0107 and TTGV (Turkish Technology Development Foundation) project no. 86.

REFERENCES

Aseev, M., Surin, V., Babodev, K., Gornostaeva, N., Kuznetzova, T., Kascheeva, T., Ivaschenko, T., Solovyev, G., Mikhailov, A., Lebedev,

V., Papaian, L., Andreeva, T., Amoashyi, D.M., Vakharlovsky, V. & Baranov, V.S. (1994) Allele frequencies and molecular diagnosis in hemophilia A and B patients from Russia and from some Asian republics of the former U.S.S.R. *Prenatal Diagnosis*, **14,** 513–522.

Beidler, J.L., Hilliard, P.R. & Rill, R.L. (1982) Ultrasensitive staining of nucleic acids with silver. *Analytical Biochemistry*, **126,** 374–380.

Broker-Vriends, A.H.J.T., Bakker, E., Kanhai, H.H., van Ommen, G.J.B., Reistma, P.H., Van de Kamp, J.J.P. & Briet, E. (1992) The contribution of DNA analysis to carrier detection and prenatal diagnosis of hemophilia A and B. *Annals of Hematology*, **64,** 2–11.

Çağlayan, S.H., Gökmen, Y., Altay, Ç., Aktuğlu, G., Gürgey, A., Soysal, T., Apak, M., Beksaç, S., Aydınlı, K. & Kırdar, B. (1995a) Carrier determination and prenatal diagnosis of hemophilia A by DNA analysis in Turkish families. *Turkish Journal of Medical Sciences*, **25,** 177–182.

Çağlayan, S.H., Gökmen, Y. & Kırdar, B. (1995b) Polymorphisms associated with the FVIII and FIX genes in the Turkish population. *Hemophilia*, **1,** 184–189.

Gitschier, J., Wood, W.I., Goralka, T.M., Wion, K.L., Chen, E.Y., Eaton, D.H., Vehar, G.A., Capon, D.J. & Lawn, R.M. (1984) Characterization of the human factor VIII gene. *Nature*, **312,** 326–330.

Goodeve, A.C., Tagariello, G., Chuansumrit, A., Preston, F.E. & Peake, I.R. (1996) A rapid and cost effective method for analysis of dinucleotide repeat polymorphisms in the factor VIII gene. *Blood Coagulation and Fibrinolysis*, **7,** 672–677.

Green, P.M., Montandon, A.J., Bentley, D.R. & Gianelli, F. (1991) Genetics and molecular biology of hemophilias A and B. *Blood Coagulation and Fibrinolysis*, **2,** 539–565.

Hoyer, L.W. (1994) Hemophilia A. *New England Journal of Medicine*, **330,** 38–47.

Lalloz, M.R., Schwaab, R., McVey, J.H., Michaelides, K. & Tuddenham, E.G. (1994) Haemophilia A diagnosis by simultaneous analysis of two variable dinucleotide tandem repeats within the factor VIII gene. *British Journal of Haematology*, **86,** 804–809.

Lalloz, M.R.A., McVey, J.H., Pattinson, J.H. & Tuddenham, E.G.D. (1991) Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet*, **388,** 207–212.

Miller, M., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, **16,** 1215.

Peake, I.R., Lillicrap, D.P., Boulyjenkov, V., Briet, E., Chan, V., Ginter, E.R., Kraus, E.M., Ljung, R., Manucci, P.M., Nicloaides, K. & Tuddenham, E.G.D. (1993) Hemophilia: strategies for carrier detection and prenatal diagnosis. *Bulletin of the World Health Organization*, **77,** 429–458.

Stoflet, E.S., Koeberl, D.D., Sarkar, G. & Sommer, S.S. (1988) Genomic amplification with transcript sequencing. *Science*, **239,** 491–494.

Windsor, S., Taylor, S.A.M. & Lillicrap, D. (1994) Multiplex analysis of two intragenic microsatellite repeat polymorphisms in the genetic diagnosis of haemophilia A. *British Journal of Haematology*, **86,** 810–815.

Yip, B., Chan, V. & Chan, T.K. (1994) Intragenic dinucleotide repeats in the factor VIIII gene for the diagnosis of haemophilia A. *British Journal of Haematology*, **88,** 889–891.