Haemophilia

Haemophilia (2011), **17**, e936–e943 DOI: 10.1111/j.1365-2516.2011.02592.x

ORIGINAL ARTICLE *Genetics*

Analysis of five polymorphic DNA markers for indirect genetic diagnosis of haemophilia A in the Brazilian population

J. D. MASSARO,* C. E. V. WIEZEL,* Y. C. N. MUNIZ,* E. M. REGO,† L. C. O. DE OLIVEIRA,† C. T. MENDES-JUNIOR‡ and A. L. SIMÕES*

*Departments of *Genetics and †Medical Clinic, School of Medicine of Ribeirão Preto, University of São Paulo; and ‡Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, SP, Brazil*

Summary. Hemophilia A is an X-linked, inherited, bleeding disorder caused by the partial or total inactivity of the coagulation factor VIII (FVIII). Due to difficulties in the direct recognition of the disease-associated mutation in the *F8* gene, indirect diagnosis using polymorphic markers located inside or close to the gene is used as an alternative for determining the segregation of the mutant gene within families and thus for detecting carrier individuals and/or assisting in prenatal diagnosis. This study characterizes the allelic and haplotype frequencies, genetic diversity, population differentiation and linkage disequilibrium of five microsatellites (*F8*Int1, *F8*Int13, *F8*Int22, *F8*Int25.3 and IKBKG) in samples of healthy individuals from São Paulo, Rio Grande do Sul and Pernambuco and of patients from São Paulo with haemophilia A to determine the degree of informativeness of these microsatellites for diagnostic purposes. The interpopulational diversity parameters highlight the differences among the analyzed population samples. Regional differences in allelic frequencies must be taken into account when conducting indirect diagnosis of haemophilia A. With the exception of IKBKG, all of the microsatellites presented high heterozygosity levels. Using the markers described, diagnosis was possible in 10 of 11 families. The *F8*Int22, *F8*Int1, *F8*Int13, *F8*Int25.3 and IKBKG microsatellites were informative in seven, six, five and two of the cases, respectively, demonstrating the effectiveness of using these microsatellites in prenatal diagnosis and in carrier identification in the Brazilian population.

Keywords: Brazilian populations, *F8*, haemophilia A, haplotypes, indirect diagnosis, microsatellites

Introduction

Haemophilia A (HEMA; OMIM 306700) is a recessive disease caused by the partial or total inactivity of the coagulation factor VIII (FVIII), a plasma glycoprotein coded by a 186-kb gene with 26 *exons* located at the Xq28 position [1]. One-third of haemophilia A cases are caused by new mutations, that is, the gene was not inherited from the patient's mothers [2]. The high number of new mutations (about 1,000 have been described so far) identified as causing haemophilia A [3] is attributed to the size of the gene, to the existence of a CpG island inside the gene and to the gene's peculiar organization (the presence of copies of the gene *F8a* in the interior and around the *F8* gene, which facilitates the occurrence of inversions) [4]. In cases of severe HEMA, the most frequent mutations are inversions in intron 22 (50%) and intron 1 (5%) [5,6]. All other mutations have extremely low frequencies [7].

Therefore, it is difficult to directly determine the mutation in each particular case. In these circumstances, the use of a haplotype formed by polymorphic markers located near or within the gene is a good alternative for determining the segregation of the mutant gene within the family being studied [8] despite the fact that this method is hardly applicable to the *de novo* mutation cases. This method is useful even when direct diagnosis is available. Therefore, more knowledge of population-specific parameters of useful markers linked to the *F8* gene is needed. The risks of recombination of the external markers limit the use of restriction fragment length polymorphisms in tracking the defective allele.

Correspondence: Aguinaldo L. Simões, PhD, Department of Genetics, School of Medicine of Ribeirão Preto, University of São Paulo, Avenida: Bandeirantes, 3900, 14049-900, Ribeirão Preto, São Paulo, Brazil.
Tel./fax: +55 16 3602 3050
e-mail: alsimoes@usp.br

Accepted after revision 11 May 2011

Additional informative intragenic markers such as short tandem repeats (STRs) in DNA would thus increase the accuracy of analyses [9].

Considering that the allele frequencies of four intronic microsatellites (*F8*Int1, *F8*Int13, *F8*Int22 and *F8*Int25.3) and an extragenic microsatellite (IKBKG) still remain to be determined, three samples of urban Brazilian populations and a group of haemophilia A patients were analyzed. The intragenic markers proved to be useful in the indirect diagnosis of carriers.

Materials and methods

Samples

Blood samples or buccal swabs were obtained from individuals from urban Brazilian populations (Fig. 1): São Paulo (SP), 160 unrelated individuals (80 men and 80 women) from cases of paternity investigation at the Hospital of the School of Medicine of Ribeirão Preto, University of São Paulo (HCFMRP-USP); Pernambuco (PE), 178 blood donors (150 men and 28 women) from Hemocentro de Pernambuco (HEMOPE) from the city of Recife; Rio Grande do Sul (RS), 154 individuals (105 men and 49 women) from the city of Porto Alegre; and haemophilia A patients (HA), 73 patients of the HEMOCENTRO of HCFMRP-USP (26 relatives from 11 patients were collected for family analyses). Of these 73 HA patients, 42 were classified as severe haemophilia cases, 14 as moderate and 17 as mild. These three groups were defined, respectively, by an FVIII coagulant activity of <1%, 1–5% and 5–40% [10]. This research was approved by the National Ethics Research Committee (Process HCRP-FMRP, USP nº 6115/2007).

Laboratory analysis

Allelic and haplotype frequencies of four intronic dinucleotide microsatellites in the *F8* gene (*F8*Int1 $(AC)_n$, *F8*Int13$(AC)_n$, *F8*Int22$(GT)_n$ and *F8*Int 25.3$(TG)_n$), as designated by Machado *et al.* [11], and one extragenic microsatellite (IKBKG) located in Intron 1 of the IKBKG gene (inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma) (Fig. 2) were estimated for each sample (Table 1). DNA extraction, loci, primers and PCR conditions have been previously described [8,12–16].

Half of the already described alleles and each of the new alleles of each *locus* were cloned and sequenced using the BigDye Terminator v3.0 Ready Reaction Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA) and an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The PCR was optimized for both the cloned inserted repeats and the genomic DNA of individuals. The amplified products were separated using electrophoresis on 12% denaturing (*F8*Int13, *F8*Int25.3 and IKBKG) or 10% non-denaturing (*F8*Int1 and *F8*Int22) polyacrylamide gels stained with silver nitrate. The cloned alleles were used as size reference for genotyping the population to determine the length of the dinucleotide repeats in tested individuals. Samples from individuals typed as above were loaded every 3–5 lanes as a reference for typing the other samples (Fig. 3).

Statistical analysis

The allelic and genotypic frequencies, the Hardy–Weinberg equilibrium, the exact tests of population differentiation and the linkage disequilibrium between pairs of microsatellites were estimated using GENEPOP software version 3.4 [17]. The haplotype diversity and

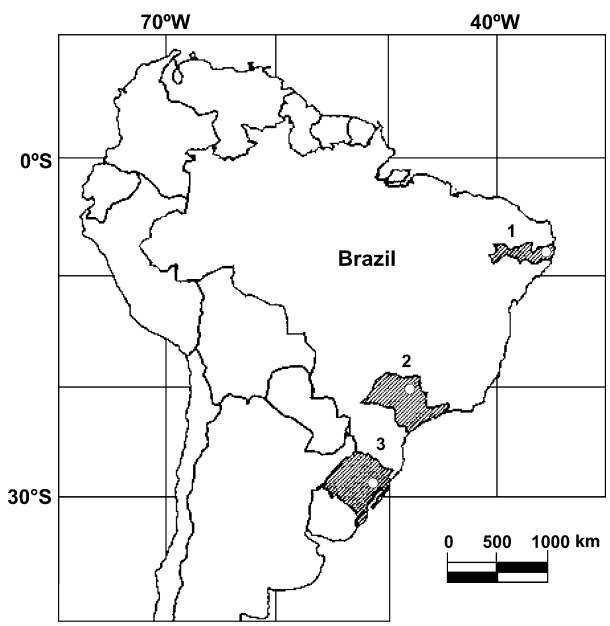

Fig. 1. Map indicating the location of the samples. (1) Recife/Pernambuco State, (2) Ribeirão Preto/São Paulo State and (3) Porto Alegre/Rio Grande do Sul State.

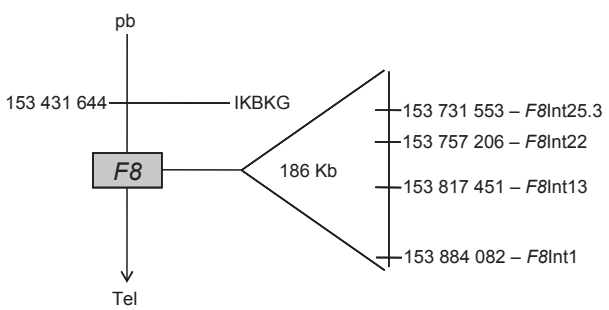

Fig. 2. Physical positions in chromosome X of the IKBKG, *F8*Int25.3, *F8*Int22, *F8*Int13 and *F8*Int1 microsatellites. The numbers correspond to pairs of bases from the Xp-Tel using the NC_000023 reference sequence.

Table 1. Allelic frequencies (%) of the *F8*Int1, *F8*Int13, *F8*Int22, *F8*Int25.3 and IKBKG microsatellites in samples of São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE) and Haemophilia A patients (HA). n = number of X chromosomes; H_k = Expected heterozygosity.

STRs/POPS	Alleles										n	H_k
*F8*Int1	**10**	**12**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**		
SP		0.4	0.4	0.4	1.7	51.5	25.5	11.7	6.3	2.1	239	0.651
RS		3.0		0.5	6.4	61.5	24.1	3.0	1.5		201	0.558
PE	0.5	0.5	0.5	1.0	9.2	50.0	21.8	6.8	8.7	1.0	206	0.682
HA					9.6	63.0	16.4	4.1	5.5	1.4	73	0.562
*F8*Int13	**16**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**		
SP		0.8	2.9	42.1	24.2	12.5	9.2	6.2	1.7	0.4	240	0.735
RS	1.0		5.9	54.5	27.3	5.9	5.4				203	0.618
PE		1.9	6.8	46.2	18.9	9.7	7.3	8.7	0.5		206	0.724
HA		1.4	8.2	45.2	23.3	5.5	9.6	6.8			73	0.718
*F8*Int22	**18**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**			
SP		2.5	6.7	8.3	37.9	40.4	3.8		0.4		240	0.680
RS			0.5	4.5	35.8	54.7	3.0	1.0	0.5		199	0.570
PE	0.5	1.0	2.4	10.2	36.4	45.1	3.9	0.5			206	0.651
HA			4.1	2.7	23.3	64.4	4.1	1.4			73	0.527
*F8*Int25.3	**12**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**			
SP		0.4	7.6	4.7	63.1	10.2	12.3	1.3	0.4		236	0.568
RS	0.5		2.0	2.5	65.7	11.9	14.9	2.5			200	0.530
PE			13.1	1.9	59.2	11.7	10.7	2.4	1.0		206	0.606
HA			8.2	4.1	68.5	2.7	13.7	1.4	1.4		73	0.502
IKBKG	**25**	**27**	**28**	**29**	**30**	**33**						
SP		0.8	6.2	86.3	6.7						240	0.247
RS		0.5	6.8	87.8	4.4	0.5					203	0.223
PE	1.0	1.0	5.9	91.1	1.0						205	0.166
HA		1.4	6.8	89.1	2.7						73	0.201

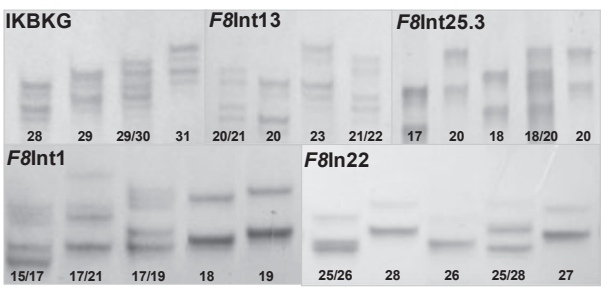

Fig. 3. Patterns of the electrophoretic mobility displayed by each microsatellite. Samples of the first and last lanes are from individuals whose phenotype was determined by cloning and sequencing. Numbers refer to the phenotype of each sample and corresponds to the number of repetitions.

the analysis of molecular variance (AMOVA) were calculated with the ARLEQUIN software [18]. The haplotypes were designated by listing the repeat numbers separated by bars in the order *F8*Int1/*F8*Int13/*F8*Int22/*F8*Int25.3/IKBKG. The genetic diversity with the standard error was calculated for each sample using DISPAN software [19]. Other parameters of diversity (H_k, $H_{T'}$ and $G_{ST'}$) were estimated using Fstat 2.8 software [20], whereas F-statistics were calculated according to Weir and Cockerham (1984) using the software GDA [21]. The bootstrap procedure with 1000 replicates was performed to test whether the values of F_{ST} differed significantly from zero. If the confidence intervals of 95% and 99% did not include the zero, the estimate was considered significantly different from zero with α = 5% or 1% respectively.

Results

The allelic frequencies and heterozygosities of the populations were analyzed, and are listed in Table 1. Some alleles occurred only once: *F8*Int1**10* in PE, *F8*Int*13*26* in SP, *F8*Int*13*16* in RS, *F8*Int*22*18* in PE, *F8*Int*25.3*12* in RS, IKBKG**25* in PE and IKBKG**33* in RS. The allelic frequencies of IKBKG and the *F8*Int*1*10*, *F8*Int*1*21* and *F8*Int*25.3*12* alleles are described here for the first time. In the SP, RS and HA female populations, all of the microsatellites exhibited the genotypic proportions predicted by the Hardy–Weinberg equilibrium. In the PE sample, all of the microsatellites except *F8*Int22 presented disequilibrium as determined by a deficit of heterozygotes.

Considering the five microsatellites and each population sample, the average H_S values were similar for men and women, grouped or not. The IKBKG microsatellite presented smaller values of total genetic diversity ($H_{T'}$ = 0,209) compared with the other four microsatellites, the $H_{T'}$ values of which of varied between 0.567 and 0.699. The $H_{T'}$ value was not significantly different between men (54.9%) and women (55.5%).

The allelic differentiation (Table 2), tested using Fisher's exact test, for the five markers in the pairwise comparisons of the four samples showed that of the 30 possible comparisons among the male fraction, 13 presented significant differences, whereas of the 15 possible comparisons in the female fraction, six comparisons were significantly different. This was a slightly larger difference than that obtained when considering the genotypic differentiation, in which only four of 15

Table 2. Genic and genotypic differentiation between possible pairs of the samples of São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE) and Haemophilia A patients (HA) based in the microsatellites *F8*Int1, *F8*Int13, *F8*Int22, *F8*Int25.3 and IKBKG.

	*F8*Int1	*F8*Int13	*F8*Int22	*F8*Int25.3	IKBKG	All STRs
Male/Female						
SP × RS	*	*	*	0.067	0.685	*
SP × PE	†	0.196	0.236	0.252	†	†
RS × PE	*	*	0.055	*	0.093	*
Male						
SP × RS	†	*	*	*	*	*
SP × PE	*	0.103	0.670	0.940	*	†
RS × PE	*	*	0.350	†	0.084	*
SP × HA	*	0.325	*	0.252	0.213	†
RS × HA	0.097	0.092	0.375	*	0.334	†
PE × HA	0.922	0.652	0.236	0.270	0.715	0.707
Female (genic)						
SP × RS	†	†	†	0.411	0.389	*
SP × PE	0.121	0.805	0.184	0.255	0.199	0.173
RS × PE	†	†	0.080	0.121	†	*
Female (genotypic)						
SP × RS	†	†	†	0.426	0.555	*
SP × PE	0.196	0.802	0.089	0.406	0.077	0.116
RS × PE	0.054	0.131	0.149	0.288	†	†

*$P \leq 0.001$.
†$0.001 < P < 0.05$.

comparisons presented significant differences. Considering all STRs, only three of the 15 possible comparisons were significantly different. Based on the five analyzed microsatellites, the pairwise *FST* values of the population showed significant differences between male patients in the pairs SP/RS, RS/HA, RS/PE and SP/HA and between female patients in the pairs SP/HA and RS/PE.

Considering men and women separately, the association analyses between all of the intronic microsatellite pairs revealed highly significant *P*-values. The IKBKG marker, the only extragenic microsatellite, did not show linkage disequilibrium with the others. A total of 163 different haplotypes (Table 3) were found in male fraction of the samples described here: 82 in PE sample, 48 in RS, 46 in SP and 45 in HA. The haplotype 17/20/26/18/29 was the most frequent in the four population samples. Moreover, 32 of the 163 different haplotypes were shared between samples, whereas 58 occurred exclusively in PE, 26 occurred in RS, 24 occurred in SP and 23 occurred in HA. The AMOVA showed that almost total diversity (99.25%) was found within the samples ($P = 0.195$).

Eleven of the 73 haemophilic patients presented relatives for pedigree analysis. It was possible to identify the haplotype that followed the mutation in 10 of these families. In the unique case in which it was not possible to identify the haplotype that followed the mutation, the patient's mother was homozygous for the most frequent haplotype (Fig. 4a). In another case (Fig. 4b), there was clearly a mutation at *F8*Int1; therefore, because we had a small pedigree with a single affected patient, it was not possible to dismiss the possibility that a new mutation caused the haemophilia in this particular patient. However, it was possible to rule out the condition of carrier for the sister of the affected. In a third case, the alleged paternity was excluded, but this fact did not interfere with the analysis. The microsatellites *F8*Int22, *F8*Int1, *F8*Int13, *F8*Int25.3 and IKBKG were informative in seven (63.6%), six (54.5%), six (54.5%), five (45.4%) and two (18.2%) cases respectively. The haplotype 17/20/26/18/29, the most frequent in the four population samples, occurred in four of the families studied.

Discussion

The occurrence of the *F8*Int*1*10*, *F8*Int*1*21* and *F8*Int*25.3*12* alleles, which were described for the first time here, was confirmed using cloning and sequencing procedures. The presence of exclusive alleles in a single population can be explained by the low frequency of these alleles (Table 4). The allelic frequencies of IKBKG, which had been briefly discussed by Fang *et al.* (2006) [22], were determined in detail here for the first time (Table 1). The low heterozygosity ($H_{T'} = 0.209$) suggests that this marker may not be useful for indirect diagnosis.

All of the microsatellites except *F8*Int22 exhibited significant deviation from the Hardy–Weinberg prediction due to a deficit of heterozygotes in the PE sample. Genotyping errors can be ruled out because genotyping errors would lead to the same disequilibrium being present in the other samples. Previous studies [23–25] in the same region (PE) did not show any special characteristics that could explain the disequilibrium. The disequilibrium is probably due to the small size of the female subsample ($n = 28$) on which our computations were based.

The weighted allele frequencies obtained for each continental group (European, Africans, Asiatics and Americans) were used to compare the present study results (Table 4). It is noteworthy that in *F8*Int1, all populations (Spanish, Korean, Indian, Chinese) studied

Table 3. Absolute frequency of the five microsatellites haplotypes of chromosome X observed in men in samples of São Paulo (SP), Rio Grande do Sul (RS), Pernambuco (PE) and Haemophilia A patients (HA). The gray boxes indicate exclusive haplotypes were not unique.

	STRs/Alleles						Number of X chromosomes			
Haplotype	*F8*Int1	*F8*Int13	*F8*Int22	*F8*Int25.3	IKBKG	***n***	SP	RS	PE	HA
1	17	20	26	18	29	**90**	15	28	32	15
2	17	21	26	18	29	**20**	3	6	4	7
3	17	19	26	18	29	**14**	1	4	6	3
4	16	20	26	18	29	**11**		3	6	2
5	17	21	25	20	29	**11**	3	4	3	1
6	17	20	26	19	29	**9**	2	1	5	1
7	18	21	25	20	29	**9**	1	6	1	1
8	20	24	25	16	29	**9**	1		6	2
9	18	20	26	18	29	**7**	1	2	3	1
10	18	23	25	18	29	**7**	3	3		1
11	17	21	25	19	29	**6**	3	1	2	
12	17	20	25	18	29	**5**		2	1	2
13	17	22	25	19	29	**5**	1	3		1
14	18	20	25	20	29	**5**	2	1	1	1
15	17	20	26	18	28	**4**	2	1		1
16	18	21	25	21	29	**4**		2	1	1
17	18	21	26	20	29	**4**	1		1	2
18	19	22	24	18	29	**4**	3	1		
19	17	18	26	18	29	**3**			2	1
20	17	20	26	18	30	**3**	2	1		
21	17	22	25	16	29	**3**	2		1	
22	18	20	25	18	29	**3**			3	
23	18	20	26	20	29	**3**			1	2
24	18	21	26	18	29	**3**	1	1	1	
25	18	22	25	20	29	**3**			3	
26	19	23	25	18	29	**3**	3			
27	12	20	26	18	29	**2**		2		
28	16	20	26	19	29	**2**		1	1	
29	17	20	26	20	29	**2**			2	
30	17	20	27	18	29	**2**				2
31	17	20	28	18	29	**2**			1	1
32	17	21	26	19	29	**2**			2	
33	17	23	25	16	29	**2**			1	1
34	18	21	25	18	29	**2**	1	1		
35	18	22	24	18	29	**2**			2	
36	18	23	26	18	29	**2**		1		1
37	19	21	24	18	29	**2**			1	1
38	19	21	25	21	29	**2**	1	1		
39	19	24	25	16	29	**2**	1		1	
40	20	23	25	16	29	**2**			2	
41	20	23	25	18	29	**2**	2			
42	21	24	25	16	29	**2**	1		1	
Unique haplotypes						**121**	22	25	52	22
Number of haplotypes						**163**	46	48	82	45
Number of shared haplotypes						**32**	22	22	24	22
Exclusive haplotypes						**131**	24	26	58	23
Number of individuals						**401**	78	101	149	73
Haplotype Diversity (h^a)							0.9570 ± 0.0158	0.9143 ± 0.0230	0.9479 ± 0.0137	0.9486 ± 0.0182

so far were different from each other in relation to the most frequent allele [16,14,15,17, respectively, Table 4], and allele 17 is the most frequent in the populations described here. It is difficult to explain this variability because even the more likely hypothesis, methodological differences in genotyping, is not sufficient to explain it.

In the European and Asian populations, the average frequencies of the *F8*Int13 alleles were centered at 20 and 21 repeat alleles. However, discrepant samples from New Zealand [26], China [27] and Indian [28] populations have been reported in which the most frequent alleles were 17 and 18, 25 and 26 and 22–24 respectively. Although allele 22 was the most frequent in the only sample of Africans reported so far (Table 4) [29], alleles 19–23 also have high frequency. The allelic distributions in Americans diverge in the three previously described samples, including a Brazilian one [30]. This makes it difficult to compare the present data, in which alleles 20 and 21 are the most frequent, with those reported in the literature.

In *F8*Int22, Europeans and Asians present distributions centered around alleles 25 and 26. Although the New Zealand sample (Table 4) is clearly different from the other samples, we do not know the precise origin of the examined patients [26]. The sample of subcontinent

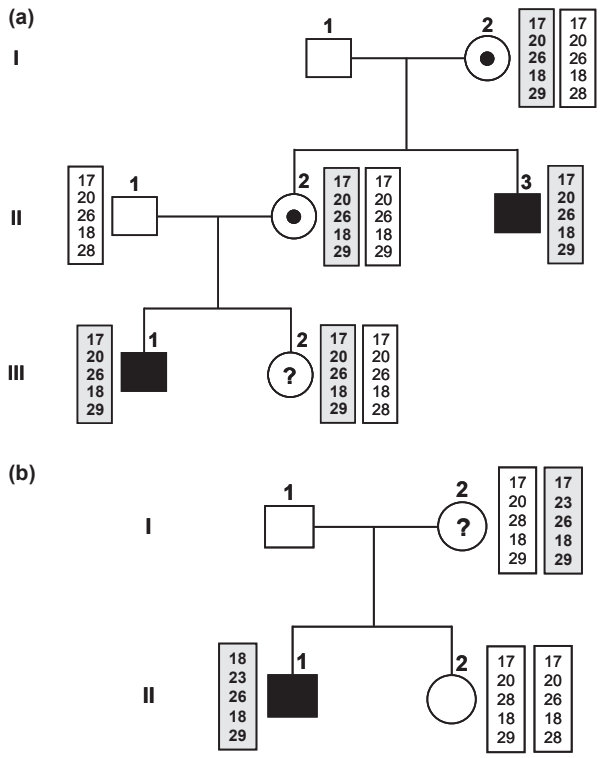

Fig. 4. Pedigree and haplotype analysis of the *F8*Int1, *F8*Int13, *F8*Int22, *F8*Int25.3 and IKBKG microsattelites in two Haemophilic A families. (a) illustrates the only one among the 11 families in which the markers were not informative; (b) The markers were informative, despite the mutation in *F8*Int1 at II-1 patient. Gray bars represent haplotype linked to mutation.

Indians [28], which diverged from the Asian average in *F8*Int13, also diverged from the reported Asian frequencies of *F8*Int22. The only sample of Africans described so far presents the 24 and 25 alleles as the most frequent ones (Table 4). These alleles are also more frequent in Canadians and Mexicans, whereas in the São Paulo sample, in spite of the remarkable frequencies of alleles 25 and 26, the most frequent allele is 27. In the samples of the present study and in the few samples previously analyzed for the *F8*Int25.3 microsatellite, allele 18 occurred at the highest frequency.

The utility of one polymorphism for prenatal disease diagnosis and/or for carrier diagnosis is directly proportional to its expected heterozygosity levels (H_K) in the population under study. For the *F8*Int1 microsatellite, the H_K values ranged from 0.56 (Spanish) to 0.66 (Asian). The highest H_K values were for *F8*Int13 for Americans (0.81), and only the Asians (0.58) showed a lesser value. The *F8*Int22 and the *F8*Int25.3 microsatellites presented H_K values ranging from 0.63 (African) to 0.71 (American) and from 0.38 (Asian) to 0.57 (present study) respectively. These differences between worldwide populations demonstrate the importance of regional surveys of allelic frequencies.

The Brazilian population is the result of interethnic crosses of European (mainly Portuguese), African and Amerindian populations, but the proportions of these components vary among different regions. The African and Amerindian contribution to the southern Brazilian population is small [31,32]. The genic and genotypic differentiation (Table 2) observed here could be due to the different ethnic composition and allelic frequency differences of the founder populations. The sampling of patients analyzed herein is composed mostly of individuals from several regions of the São Paulo State. Therefore, some degree of differentiation between HA and the other samples could be expected.

The exact test of population differentiation based on allele frequencies (Table 2) and pairwise F_{ST} analysis showed an increased number of significant differences when men were considered in the comparison. The differences in the inter-populational diversity parameters are in agreement with the analysis of molecular variance. These results highlight the differences among the analyzed population samples, which could be the consequence of differences in the contribution of founder populations to each of the regional samples [31,32]. Thus, the regional differences in allelic frequencies must be taken into account when conducting an indirect diagnosis of haemophilia A.

The population differences were less pronounced when considering the IKBKG microsatellite, which could be explained by the low heterozygosity of this microsatellite. In addition, it is the only microsatellite that did not show linkage disequilibrium with the other microsatellites analyzed. This is probably due to IKBKG being the most distant from the others, which would increase the possibility of recombination.

The comparison of haplotypes of the present data with those in the literature is possible only for the *F8*Int13 and *F8*Int22 microsatellites. The most frequent haplotype in the present data was 20/26 (35.24%). The most frequent haplotypes (*F8*Int19 *Hin*dIII, *F8*Int13, *F8*Int22 and *F8*Int22 *Xba*I) observed in Turkish patients and healthy individuals are formed by the alleles –/20/26/+ and +/21/25/– respectively [33].

The haplotype distribution was similar in South-African black patients and healthy individuals (with or without the Intron 22 inversion). The frequency of the 20/26 (*F8*Int13/*F8*Int22) haplotype in South-African Caucasoid patients without the Intron 22 inversion was higher than in healthy individuals ($P < 0.01$). The finding of a low inversion frequency in this group together with the disproportional representation of one haplotype on chromosomes without inversions is suggestive of the presence of a founder haemophilia A mutation [29].

Due to convergent evolution in microsatellites, the interpretation of haplotype frequencies across populations should be conducted with caution. Two or more populations can share haplotypes at a high frequency because of convergent mutations, drift, gene flow or a common origin [34]. For this reason, the researcher

Table 4. Comparison of allelic frequencies and heterozygosities for *F8*Int1, *F8*Int13, *F8*Int22 and *F8*Int25.3 microsatellites in population samples of different continents. Figures of present study samples and those based in more than one bibliography reference represent the weight mean. *n* = number of X chromosome; H_k = expected heterozygosity.

STRS/population	Alleles														*n*	H_K
*F8*Int1	10	12	13	14	15	16	17	18	19	20	21					
Spanish [36]					1.4	62.7	22	7.0	5.7	1.4					282	56.0
Korean [8]		0.1	0.5	79.2	19.2	1.0									1000	33.6
Indian [28]			1.5	1.5	47.9	37.3	10.3	1.5							126	46.0
Chinese [37]				0.2	0.5	0.9	75.9	21.1	0.9	0.5					440	37.9
Present study	0.1	1.1		0.3	0.6	6.0	55.0	23.0	7.1	5.6	1.1				719	63.2
*F8*Int13	14	16	17	18	19	20	21	22	23	24	25	26	27	28		
New Zealand [26]		2.0	48.0	30.0	5.0	15.0									100	72.0
European [38–40]		0.5	11.7	7.3	2.9	54.5	24.5	8.9	7.9	1.3					311	62.7
Asian [8,9,37,40–44]	0.03		0.1	0.4	3.6	63.7	22.3	3.5	4.4	0.4	1.4				2490	54
Chinese [27]								1.0	1.0	9.0	53.0	32.0	3.0	1.0	100	61.0
Indian [28]						2.0	4.8	34.5	22.3	29.5	2.7	1.4	1.4	1.4	126	46.0
African Negroids [29]				2.0	13.0	11.0	13.0	29.0	19.0	9.0		4.0			47	82.0
Canadian [45]			1.7	1.2	4.0	7.0	25.0	47.7	13.0	0.4					243	68.0
Mexican [46]					2.6	9.3	16.0	28.8	14.0	14.0	10.0	5.3			166	83.0
São Paulo [30]							7.8	3.1	4.7	21.9	14.1	25.0	17.2	6.2	64	83.0
Present study		0.3		0.9	5.4	32.0	31.0	15.1	7.8	6.7	0.7	0.1			722	75.5
*F8*Int22	18	19	20	21	22	23	24	25	26	27	28	29				
New Zealand [26]			3.0	30.0	60.0	3.0	4.0								100	62.0
European [14,40]						0.7	2.4	35.3	57.3	3.6	0.7				417	54.5
Asian [9,27,37,40–42,44,47]			0.07			0.5	5.7	26.7	60.1	5.6	1.1	0,2			1355	56.0
Indian [28]	38.3	51.7	2.9	7.1											126	42.8
Korean [8]						0.1	20.2	71.4	7.9	0.2	0.2				1000	44.3
African Negroids [29]	2.0				2.0	11.0	23.0	47.0	9.0	6.0					47	70.0
American [30,45,46]				0.5	1.1	9.3	41.5	32	7.5	7.6	0.5				361	70.6
Present Study	0.1				1.1	3.5	7.2	35.4	48.2	3.6	0.6	0.3			718	63.5
*F8*Int25.3	9	12	14	15	16	17	18	19	20	21	22	23				
Spanish [15]					3.2		75.8	9.0	10.5	1.5					100	41.0
Asian [8,37]	0.1		0.1	2.3	2.8	1.0	78.1	6.4	7.9	0.8	0.4	0.1			1440	37.7
Present study		0.1		0.1	7.7	3.2	63.3	10.3	12.7	2.0	0.6				715	56.6

must be careful when making conclusions based on samplings of restricted areas. However, the existence of significant differences between populations in the allele frequencies suggests that local patterns of variability of these microsatellites should be explored to evaluate their forensic usefulness. The IKBKG marker presented a low rate of heterozygosity, contributed to family analysis in just two (18.2%) cases and did not show linkage disequilibrium with the intronic markers. Therefore, situations in which the diagnosis is based exclusively on data from this microsatellite should be carefully considered.

The knowledge of the linkage phase of the five microsatellites altered the genetic counseling in 10 of 11 cases in our sample. This informativeness is similar to that described by Soares *et al.* 2001 [30], who reported 100% cumulative reliability using four informative polymorphisms, but is higher than that obtained by de Carvalho *et al.* 2007 [35] using three polymorphisms. Although diagnosis of haemophilia A from linkage analysis has limitations, the results demonstrate the effectiveness of the use the microsatellites descried here in prenatal diagnosis and in the identification of carriers in the Brazilian population.

Acknowledgements

We thank Maria do Carmo Tomitão Canas and Ana Lúcia Pimentel for technical support.

Disclosures

The authors stated that they had no interests which might be perceived as posing a conflict or bias.

References

1 Gitschier J, Wood WI, Goralka TM *et al.* Characterization of the human factor VIII gene. *Nature* 1984; **312**: 326–30.
2 Larner AJ. The molecular pathology of haemophilia. *Q J Med* 1987; **63**: 473–91.
3 Kemball-Cook G. The Hemophilia A Mutation, Structures, Test and Resource Site: HAMSTeRS, 2010. Available at: http://hadb.org.uk. Accessed May 28, 2010.
4 Rodgers GM, Greenberg CS. Inherited coagulation disorders. In: Lee GR, Foerster J, Lukens J, Paraskevas F, Greer JP, Rodgers GM eds. *Wintrobe's Clinical Hematology, 10th*. Baltimore, MD: Williams & Wilkins, 1999: 1682–732.
5 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.
6 Bagnall RD, Waseem N, Green PM, Giannelli F. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood* 2002; **99**: 168–74.
7 Pandey GS, Mittal B. Molecular diagnosis in haemophilia A. *J Postgrad Med* 2001; **47**: 274–80.
8 Kim JW, Park SY, Kim JM, Kim DJ, Ryu HM. Identification of new dinucleotide-repeat polymorphisms in factor VIII gene using

fluorescent PCR. *Haemophilia* 2005; **11**: 38–42.
9 Rabbani B, Rezaeian A, Khanahmad H, Bagheri R, Kamali E, Zeinali S. Analysing two dinucleotide repeats of FVIII gene in Iranian population. *Haemophilia* 2007; **13**: 740–4.
10 Eyster ME, Lewis JK, Shapiro SS, Gill F, Kajani M, Prager D. The Pennsylvania hemophilia program. *Am J Hematol* 1978; **9**: 277–86.
11 Machado FB, Medina-Acosta E. High-resolution combined linkage physical map of short tandem repeat loci on human chromosome band Xq28 for indirect haemophilia A carrier detection. *Haemophilia* 2009; **15**: 297–308.
12 Higuchi R. Simple and rapid preparation of samples for PCR. In: Erlich HA ed. *PCR technology – principles and applications for DNA amplification*. NY: Stockton Press, 1989: 31–8.
13 Lalloz MRA, Mcvey JH, Pattinson JK, Tuddenham EGD. Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet* 1991; **338**: 207–11.
14 Lalloz MRA, Schwaab R, Mcvey JH, Michaelides K, Tuddenham EGD. Haemophilia A diagnosis by simultaneous analysis of two variable dinucleotide tandem repeats withim the factor VIII gene. *Br J Haematol* 1994; **86**: 804–9.
15 Venceslá A, Baena M, Fares TaieL, Cornet M, Baiget M, Tizzano EF. Application of intron 9 and intron 25 dinucleotide repeats of the factor VIII gene for carrier diagnosis in haemophilia A. *Haemophilia* 2008; **14**: 489–93.
16 Machado FB, Medina-Acosta E. An IKBKG, and not a G6PD, short tandem repeat marker is used in indirect diagnosis of haemophilia A. *Haemophilia* 2008; **14**: 849–50.
17 Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact test and ecumenicism. *J Hered* 1995; **86**: 248–9.
18 Scheneider S, Roessli D, Excoffier L. *Arlequin ver 2000: a software for population genetics data analysis*. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva, 2000.
19 Ota T. *DISPAN: genetic distance and phylogenetic analysis*. State College, PA: Institute of Molecular Evolutionary Genetics, Pennsylvania State University, 1993.
20 Goudet J. A program to estimate and test gene diversities and fixation indices FSTAT (version 2.8). Updated from Godet (1995), 1999.
21 Lewis PO, Zaykin D. *Genetic data analysis: computer program for the analysis of allelic data (distributed by the authors)*. Storrs, CT: Department of Ecology and Evolution, University of Connecticut, 2000.
22 Fang Y, Wang XF, Dai J, Wang HL. A rapid multifluorescent polymerase chain reaction for genetic counselling in Chinese haemophilia A families. *Haemophilia* 2006; **12**: 62–7.
23 Leboute APM. *Polimorfismos Genéticos em três amostras da população urbana brasileira*. USP, SP, Brazil: Tese de Doutorado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, 2000.
24 Mauricio-da-Silva L, Silva RS, Dellalibera E, Donadi EA. Population genetics of HPRTB, F13B, and LPL in Pernambuco, Northeast Brazil. *J Forensic Sci* 2000; **45**: 684–6.
25 Dellalibera E, Havro ML, Souza M, Kajihara K, Mauricio-da-Silva L, dos Santos Silva R. Genetic analysis of 13 STR loci in the population from the State of Pernambuco, northeast Brazil. *Forensic Sci Int* 2004; **146**: 57–9.
26 Harraway JR, Smith MP, George PM. A highly informative, multiplexed assay for the indirect detection of hemophilia A using five-linked microsatellites. *J Thromb Haemost* 2006; **4**: 587–90.
27 Yip B, Chan V, Chan TK. Intragenic dinucleotide repeats in factor VIII gene for the diagnosis of haemophilia A. *Br J Haematol* 1994; **88**: 889–91.
28 Saha A, Mukherjee S, Maulik M, Chandak GR, Indian Genome Variation Consortium, Ray K. Evaluation of genetic markers linked to hemophilia A locus: an Indian experience. *Haematologica* 2007; **92**: 1725–6.
29 Dangerfield BT, Manga P, Field SP, Hartman E, Jemkins T, Krause A. Feasibility of prenatal diagnosis and carrier detection in South African haemophilia A patients. *Br J Haematol* 1997; **97**: 558–60.
30 Soares RP, Chamone DA, Bydlowski SP. Factor VIII gene inversions and polymorphisms in Brazilian patients with haemophilia A: carrier detection and prenatal diagnosis. *Haemophilia* 2001; **7**: 299–305.
31 Ferreira LB, Mendes-Junior CT, Wiezel CEV, Luizon R, Simões AL. Genomic ancestry of a sample population from the state of São Paulo, Brazil. *Am J Hum Biol* 2006; **18**: 702–5.
32 Muniz YC, Ferreira LB, Mendes-Junior CT, Wiezel CE, Simões AL. Genomic ancestry in urban Afro-Brazilians. *Ann Hum Biol* 2008; **35**: 104–11.
33 El-Maarri O, Kavakli K, Çaglayan H. Intron 22 inversions in the Turkish haemophilia A patients: prevalence and haplotype analysis. *Haemophilia* 1999; **5**: 169–73.
34 Thomas MG, Parfitt T, Weiss DA *et al*. Y chromosomes traveling south: the cohen modal haplotype and the origins of the Lemba – the ''Black Jews of Southern Africa''. *Am J Hum Genet* 2000; **66**: 674–86.
35 de Carvalho FM, de VargasWolfgramm E, Paneto GG *et al*. Analysis of Factor VIII polymorphic markers as a means for carrier detection in Brazilian families with haemophilia A. *Haemophilia* 2007; **13**: 409–12.
36 Tizzano E, Venceslá A, Cornet M, Baena M, Baiget M. Utility of a (GT) dinucleotide repeat in intron 1 of the factor 8 gene for haemophilia A carrier diagnosis. *Haemophilia* 2005; **11**: 142–4.
37 Liang Y, Zhao Y, Yan M, Fan XP, Xiao B, Liu JZ. Prenatal diagnosis of haemophilia A in China. *Prenat Diagn* 2009; **29**: 664–7.
38 Aseev M, Surin V, Baboev K *et al*. Allele frequencies and molecular diagnosis in haemophilia A and B patients from Russia and from some Asian republics of the former U.S.S.R. *Prenat Diagn* 1994; **14**: 513–22.
39 Sánchez-García JF, Gallardo D, Ramírez L, Vidal F. Multiplex fluorescent analysis of four short tandem repeats for rapid haemophilia A molecular diagnosis. *Thromb Haemost* 2005; **94**: 1099–103.
40 Goodeve AC, Tagariello G, Chuansumrit A, Preston FE, Peake IR. A rapid and cost effective method for analysis of dinucleotide repeat polymorphisms in the factor VIII gene. *Blood Coagul Fibrinolysis* 1996; **7**: 672–7.
41 Lin SR, Chang SC, Lee CC, Shen MC, Lin SW. Genetic diagnosis of haemophilia A of Chinese origin. *Br J Haematol* 1995; **91**: 722–7.
42 Jarjanazi H, Timur AA, El-Maarri O, Caglayan SH. Analysis of the two microsatellite repeat polymorphisms of the factor VIII gene in the Turkish population. *Br J Haematol* 1998; **100**: 589–93.
43 Sawada A, Sumita C, Higasa S, Ueda M, Suehiro A, Kakishita E. Suitability of four polymorphic DNA markers for indirect genetic diagnosis of haemophilia A in Japanese subject. *Thromb Res* 2002; **105**: 271–6.
44 Jayandharan G, Shaji V, George B, Chandy M, Srivastava A. Informativeness of linkage analysis for genetic diagnosis of haemophilia A in India. *Haemophilia* 2004; **10**: 553–9.
45 Windsor S, Taylor SA, Lillicrap D. Direct detection of a common inversion mutation in the genetic diagnosis of severe hemophilia A. *Blood* 1994; **84**: 2202–5.
46 Gallegos RM, Aranda HB, Navarrete CP, Espinoza R, Gomez FS, Aranda DA. Polymorphism distribution of Int13, Int22, and St14 VNTRs in a Mexican population and their application in carrier diagnosis of hemophilia A. *Am J Hematol* 2004; **77**: 1–6.
47 Chowdhury MR, Herrmann FH, Schroder W *et al*. Factor VIII gene polymorphisms in the Asian Indian population. *Haemophilia* 2000; **6**: 625–30.