Iranian Journal of Basic Medical Sciences

ijbms.mums.ac.ir

IJBMS

Heterozygosis deficit of polymorphic markers linked to the β-globin gene cluster region in the Iranian population

Tahereh Moradi [1], Reihaneh Vallian [2], Zahra Fazeli [3], Asieh Haghighatnia [2], Sadeq Vallian [1, 2]*

[1] Division of Genetics, Department of Biology, Faculty of Science, University of Isfahan, Isfahan, Iran
[2] Molecular Genetics Department, Isfahan Medical Genetics Center, Isfahan, Iran
[3] Department of Genetics, Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Article type:
Original article

Article history:
Received: Oct 6, 2014
Accepted: Feb 16, 2015

Keywords:
β-globin gene cluster region
β-thalassemia
Iranian population
Polymorphic markers

ABSTRACT

Objective(s): Iran is considered as one of the high-prevalence areas for β-thalassemia with a rate of about 10% carrier frequency. Molecular diagnosis of the disease is performed both by direct sequencing and indirectly by the use of polymorphic markers present in the beta globin gene cluster. However, to date there is no reliable information on the application of the markers in the Iranian population. Here we report the results of an extended molecular analysis of five RFLP markers, *Xmn*I, *Hind*IIIA, *Hind*IIIG, *Rsa*I and *Hinf*I, located within the β-globin gene cluster region in four subpopulations of Iran.

Materials and Methods: A total of 552 blood samples taken from the Iranian subpopulations including Isfahan, Chaharmahal-O-Bakhtiari, Khuzestan and Hormozgan were genotyped using PCR-RFLP and sequencing. The allele frequency, the expected and observed heterozygosity, and Shannon's information index (I) of these markers were calculated.

Results: Distribution of the allele frequencies for *Xmn*I, *Hind*IIIA, *Hind*IIIG, *Rsa*I and *Hinf*I polymorphic markers did not differ significantly among the subpopulations examined. Overall observed heterozygosity ranged from 0.1706 for *Hind*IIIA to 0.4484 for *Rsa*I. The Shannon index was <1 for all the polymorphic markers in the populations studied. The data indicated that heterozygosity of these markers was low in the Iranian population.

Conclusion: The results suggested that genotyping of these markers is not informative enough once used as single markers for prenatal diagnosis and carrier detection of β-thalassemia in the Iranian population. However, haplotyping of these markers may provide more useful data in linkage analysis and prenatal diagnosis as well as carrier detections for β-thalassemia in Iranians.

► *Please cite this paper as:*
Moradi T, Vallian R, Fazeli Z, Haghighatnia A, Vallian S. Heterozygosis deficit of polymorphic markers linked to the β-globin gene cluster region in the Iranian population. Iran J Basic Med Sci 2015; 18:571-575.

Introduction

The hemoglobinopathies are the most common inherited single-gene disorders worldwide, resulting from the heterozygote advantage against malaria (1). Among the hemoglobin disorders, α-thalassemias are found with high gene frequencies in African, Asian and the Mediterranean populations, whereas β-thalassemias are particularly prevalent in Asia and the Mediterranean natives (2). The β-thalassemias are a group of anemia which affect the rate of synthesis of normal β-globin chains. β-thalassemia occurs due to the mutations in the β-globin gene cluster located on short arm of chromosome 11 (3). More than 500 mutations have been characterized in β-globin gene region associated with the β-thalassemia disease (OMIM 141900). However, it is revealed that each population has its own spectrum of mutations (4). The beta globin gene cluster region contains five expressed genes with the same transcriptional orientation and one pseudogene (ψβ) between Aγ and δ genes in the order of 5’-ε-Gγ-Aγ-ψβ-δ-β-3’ (5-6). Moreover, the cluster region contains several polymorphic markers throughout chromosome 11, which could be used as useful tools for linkage analysis and molecular diagnosis of beta thalassemia disease (7).

β-thalassemia is one of the most common public health problems in Iran, having the frequency of about 10% in the Iranian population (8). More than two million carriers of β-thalassemia live in Iran (9). Therefore, molecular diagnosis of the disease is highly requested in most molecular diagnosis centers in Iran. The disease is diagnosed both by direct sequencing and indirectly by the use of polymorphic markers present in the beta globin gene cluster.

However, to date there is no reliable information on the application of the markers in the Iranian population. In this study, the aim was to investigate

**Corresponding author:* Sadeq Vallian. Division of Genetics, Department of Biology, Faculty of Science, University of Isfahan, Isfahan, Iran. Tel/Fax: +98-31-37932456; email: svallian@sci.ui.ac.ir

Table 1. The frequency distribution of RFLP markers located within β -globin gene cluster region in four subpopulations of Iran

Population	Sample size (2n)	Allele	XmnI	HindIIIA	HindIIIG	RsaI	HinfI
Isfahan	176	-	0.7105	0.8553	0.5592	0.3947	0.2237
		+	0.2895	0.1447	0.4408	0.6053	0.7763
Chaharmahal-O-Bakhtyari	201	-	0.8564	0.8911	0.7228	0.3564	0.1584
		+	0.1436	0.1089	0.2772	0.6436	0.8416
Khuzestan	97	-	0.6915	0.9255	0.6489	0.3830	0.2766
		+	0.3085	0.0745	0.3511	0.6170	0.7234
Hormozgan	78	-	0.8571	0.8929	0.6964	0.5536	0.1071
		+	0.1429	0.1071	0.3036	0.4464	0.8929
Overall	552	-	0.7817	0.8869	0.6567	0.3948	0.1944
		+	0.2183	0.1131	0.3433	0.6052	0.8056

The presence or absence of the restriction sites are represented by (+) and (-) symbols in this table

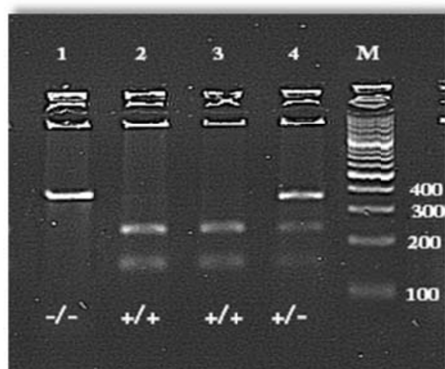

Figure 1. A typical illustration of *Hinf*I marker genotyping in four DNA samples from the Iranian population. Total genomic DNA was extracted and digested with *Hinf*I restriction enzyme. The digestion products were resolved on 2% agarose gel and visualized following ethidum bromide staining. Lane 1: homozygous -/-, lanes 2 and 3: homozygous +/+, lane 4: heterozygous +/-. M represents 100 bp DNA ladder

the population indices of five most commonly used polymorphic markers located in the beta globin gene cluster region including *Xmn*I, *Hind*IIIA, *Hind*IIIG, *Rsa*I and *Hinf*I in the β-thalassemia carrier individuals from central and south provinces of Iran.

Materials and Methods

Genomic DNA samples and genotyping

Blood samples were collected from 552 unrelated β-thalassemia carriers referred to the Isfahan Medical Genetics Center, Isfahan, Iran. Informed consent was taken from all the individuals participated in the study. Also, the ethical committee of University of Isfahan has approved the study (approval number 790205). The selected individuals were from the provinces of Isfahan (n=176), Chaharmahal-O-Bakhtyari (n=201), Khuzestan (n=97) and Hormozgan (n=78). Genomic DNA from each individual was obtained from peripheral leucocytes by salting out procedure (10). Five RFLP markers including *Xmn*I 5′ to the $^{G}\gamma$ gene, *Hind*III in the IVSII of $^{G}\gamma$ and $^{A}\gamma$ genes, *Rsa*I 5′ and *Hinf*I 3′ to the β-globin gene were genotyped using PCR-RFLP technique (11, 12). The primers and PCR-RFLP conditions are available on request.

Statistical analysis

The statistical analysis of data was performed by use of the Popgene32 software version 1.31 (available at http://www.ualberta.ca/~ fyeh/-download.htm) and PyPop program (13). The allele frequency, the expected and observed heterozygo-sity, Shannon's information index (I), Fixation indices and the Ewens-Watterson test were estimated. The D' values between possible pairings of studied markers were obtained by means of the 2LD program (14).

Results

Five polymorphic markers in the beta globin gene region including *Xmn*I, *Hind*IIIA, *Hind*IIIG, *Rsa*I and *Hinf*I were genotyped in two central populations of Iran, Isfahan and Chaharmahal-O-Bakhtyari, as well as two south populations, Khuzestan and Hormozgan (Figure 1). As shown in Table 1, allele frequency of the markers did not differ significantly among the four Iranian subpopulations. The expected and observed hetero-zygosity, Shannon's information index (I), and F_{is} value of the markers are shown in Table 2. The overall observed heterozygosity ranged from 0.1706 (*Hind*IIIA marker) to 0.4484 (*Rsa*I marker). Negative F_{is} was observed for *Hinf*I markers in Khuzestan and *Rsa*I marker in Hormozgan, indicating an excess of heterozygotes. All these polymorphic markers showed the Shannon index <1 in the studied populations,

revealing low heterozygosity of studied markers in these populations of Iran (Table 2).

Table 2. The observed and expected heterozygosity, Shannon's information index (I) and inbreeding coefficient (F_{is}) of five RFLP markers located within β -globin gene cluster region in four subpopulations of Iran

Population	RFLP marker	Obs. Het.	Exp. Het.	I	Chi-square	*P-value*	F_{is}
Isfahan	*Xmn*I	0.3684	0.4141	0.6017	0.941879	0.331795	0.1044
	*Hind*IIIA	0.2368	0.2492	0.4135	0.195192	0.658630	0.0434
	*Hind*IIIG	0.4342	0.4963	0.6861	1.204216	0.272500	0.1192
	*Rsa*I	0.4737	0.4810	0.6708	0.017856	0.893698	0.0087
	*Hinf*I	0.3158	0.3496	0.5315	0.728893	0.393242	0.0907
Chaharmahal-O-Bakhtyari	*Xmn*I	0.2475	0.2471	0.4114	0.000265	0.987015	-0.0066
	*Hind*IIIA	0.1386	0.1951	0.3442	8.822704	0.002975	0.2859
	*Hind*IIIG	0.2970	0.4027	0.5903	7.062978	0.007869	0.2588
	*Rsa*I	0.3762	0.4611	0.6513	3.458935	0.062911	0.1799
	*Hinf*I	0.2574	0.2680	0.4370	0.160687	0.688525	0.0346
Khuzestan	*Xmn*I	0.3191	0.4313	0.6179	3.269645	0.070573	0.2520
	*Hind*IIIA	0.1489	0.1393	0.2650	0.258220	0.611346	-0.0805
	*Hind*IIIG	0.3617	0.4605	0.6481	2.221162	0.136131	0.2062
	*Rsa*I	0.4681	0.4777	0.6655	0.019480	0.889000	0.0096
	*Hinf*I	0.4681	0.4045	0.5897	1.200467	0.273228	-0.1697
Hormozgan	*Xmn*I	0.2143	0.2494	0.4101	0.525135	0.468660	0.1250
	*Hind*IIIA	0.1429	0.1948	0.3405	2.351020	0.125201	0.2533
	*Hind*IIIG	0.2500	0.4305	0.6139	5.176500	0.022894	0.4087
	*Rsa*I	0.6071	0.5032	0.6874	1.238710	0.265720	-0.2284
	*Hinf*I	0.1429	0.1948	0.3405	2.351020	0.125201	0.2533
Total	*Xmn*I	0.2937	0.3419	0.5247	5.060398	0.024479	0.1395
	*Hind*IIIA	0.1706	0.2010	0.3529	5.846419	0.015609	0.1494
	*Hind*IIIG	0.3452	0.4518	0.6432	14.078405	0.000175	0.2343
	*Rsa*I	0.4484	0.4788	0.6709	1.021542	0.312154	0.0617
	*Hinf*I	0.3016	0.3139	0.4926	0.390791	0.531884	0.0373

P-value= Probability; Obs. Het. = Observed heterozygosity; Exp. Het. = Expected heterozygosity

Fixation indices F_{IS}, F_{ST}, F_{IT} were used to analyze genetic variation. As shown in Table 3, the F_{ST} varied among studied markers from 0.0064 (*Hind*IIIA) to 0.0355 (*Xmn*I) with a mean of 0.0231. The F_{IT} values ranged from 0.0134 in *Rsa*I marker to 0.2553 in *Hind*IIIG polymorphic marker. However, mean F_{IS} (0.1025) and F_{IT} (0.1232) were both positive and greater than zero. The subscripts I, S and T refer to individual, population and total population, respectively. The mean number of migrants per generation among populations (*Nm*) was reported maximum for the RFLP marker *Hind*IIIA (38.9648) and minimum for *Xmn*I marker (6.7938).

In Ewens-Watterson test of neutrality for these studied RFLP markers, the observed homozygosity (F) lied inside the limit of 95% confidence region, and the normalized deviate of the homozygosity (F_{nd}) was estimated negative for all these polymorphic markers (Table 4).

Pairwise linkage disequilibrium values were also estimated using the 2LD programs (Zhao 2004). The obtained results are illustrated in Table 5. The number of samples of Hormozgan population was insufficient, and linkage disequilibrium values were only assessed for Isfahan, Chaharmahal-O-Bakhtyari and Khuzestan populations. D' values ranged from 0.01516 (*Hind*IIIA-*Hinf*I) to 0.81098 (*Xmn*I-*Hind*IIIG) in the studied Iranian subpopulations.

Discussion

The usefulness of the polymorphic markers in linkage study and prenatal diagnosis as well as carrier detection of genetic diseases depends on the frequency and informative status of the markers in each population. As shown in Table 1, the frequency of the minor alleles exceeds 0.3 for *Hind*IIIG and *Rsa*I markers. These data indicate that these two polymorphic markers, *Hind*IIIG and *Rsa*I, have a high degree of polymorphism in the Iranian population. Significant difference between the observed and expected heterozygosity was not observed for almost any of the RFLP markers in the studied Iranian population. The observed heterozygosity was lower than the expected heterozygosity for *Hind*IIIG and

Table 3. *F*-statistics for five RFLP markers located within β -globin gene cluster region in the Iranian population

RFLP marker	F_{IS}	F_{IT}	F_{ST}	N_m
*Xmn*I	0.1350	0.1657	0.0355	6.7938
*Hind*IIIA	0.1344	0.1399	0.0064	38.9648
*Hind*IIIG	0.2422	0.2553	0.0172	14.2843
*Rsa*I	-0.0114	0.0134	0.0245	9.9656
*Hinf*I	0.0177	0.0438	0.0266	9.1389
Mean	0.1025	0.1232	0.0231	10.5792

Table 4. The Ewens-Watterson homozygosity test of neutrality for five RFLP markers located within β -globin gene cluster region in the Iranian subpopulations

RFLP marker	K	F	$\hat{F}$	F_{nd}	L95	U95
*Xmn*I	2	0.6588	0.8539	-1.1768	0.5045	0.9960
HindIIIA	2	0.7994	0.8539	-0.3287	0.5042	0.9960
HindIIIG	2	0.5491	0.8539	-1.8380	0.5053	0.9960
*Rsa*I	2	0.5221	0.8539	-2.0009	0.5045	0.9960
*Hinf*I	2	0.6867	0.8539	-1.0081	0.5035	0.9960

k: the number of alleles; F: observed homozygosity; $\hat{F}$: the expected homozygosity; L95, U95: the 95% confidence interval upper and lower limit

*Rsa*I in Hormozgan and Chaharmahal-O-Bakhtyari populations and for *Xmn*I and *Hind*IIIG in the Khuzestan population. These results indicate the presence of heterozygosis deficit of these markers in these populations. The observed heterozygosity of the RFLP markers was lower than 50% in all the studied populations, indicating that these markers are less informative in prenatal diagnosis and carrier detection of β-thalassemia. However, haplotyping of the markers could be more useful than separately genotyping them in the Iranian population.

Analysis of Shannon index indicated that the highest value was observed for *Rsa*I polymorphism in all Iranian subpopulations studied. These data revealed that *Rsa*I polymorphism was the most useful markers among the five RFLP markers investigated. The *Hind*IIIA polymorphism in all the studied populations, and *Xmn*I and *Hinf*I polymorphisms in Hormozgan and Chaharmahal-O-Bakhtyari populations showed a very low Shannon index. These results reveal that these markers have little application in carrier detection and prenatal diagnosis of β-thalassemia disease in the Iranian population.

Moreover, the studied RFLP markers did not show significant deviations from Hardy Weinberg Equilibrium (HWE) except *Hind*IIIA and *Hind*IIIG polymorphisms in Chaharmahal-O-Bakhtyari population and *Hind*IIIG polymorphism in Hormozgan population. Once the subpopulations were considered as one population, the population only deviated from HWE at *Xmn*I, *Hind*IIIA and *Hind*IIIG polymorphic markers. In the calculation of HWE, it should be noted that the selected individuals were unrelated with no consanguineous relationship (their parents were unrelated). However, whether their parental marriage was random or not, is not clear. Therefore, we cannot definitely exclude any population inbreeding and claim that all marriages in that population were random.

Genetic structure of these populations was analyzed through Wright's *F*-statistics. F_{ST} is the most suitable statistic for studying the overall level of genetic divergence (15). The data showed that mean F_{ST} was 0.0231. This indicated low level of population differentiation and high inbreeding among the Iranian subpopulations. The values of mean F_{IS} and F_{IT} were both greater than zero. This suggested the presence of a heterozygosis deficit within the studied populations. The value of *Nm* was 10.5792 (Table 3). When *Nm* is higher than 1, gene flow is able to offset the differentiation among populations caused by isolation and genetic drift (16).

In the Iranian population, F value lied inside the lower and upper limits of the 95% confidence region

Table 5. Linkage disequilibrium values for possible pairings of markers located within β-globin gene cluster region in four Iranian subpopulations

Population	D'				
		*Hind*IIIA	*Hind*IIIG	*Rsa*I	*Hinf*I
Isfahan	*Xmn*I	0.524066	0.809143	0.026556	0.331211
	*Hind*IIIA		0.691725	0.015976	0.999867
	HindIIIG			0.007165	0.218419
	*Rsa*I				0.879942
Chaharmahal-O-Bakhtyari	*Xmn*I	0.933551	0.832847	0.151087	0.111049
	*Hind*IIIA		0.672099	0.149183	0.227101
	*Hind*IIIG			0.001400	0.054807
	*Rsa*I				0.554129
Khuzestan	*Xmn*I	0.313749	0.707050	0.040465	0.365207
	*Hind*IIIA		1.000000	0.970610	0.231064
	*Hind*IIIG			0.120297	0.250281
	*Rsa*I				0.823589
Total	*Xmn*I	0.58026	0.81098	0.08158	0.27417
	*Hind*IIIA		0.70023	0.09390	0.01516
	*Hind*IIIG			0.02481	0.18810
	*Rsa*I				0.77609

of expected F values for all these RFLP markers (Table 4), signifying that these polymorphic markers in β-globin gene region were not under genetic hitchhiking and selection force operating on another locus could not change allele frequency and heterozygosity of these polymorphic markers in this population.

As shown in Table 4, the expected homozygosity ($\hat{F}$) was higher than the observed homozygosity (F) value for all these polymorphic markers in the β-globin gene cluster. The normalized deviate of the homozygosity (F_{nd}) was also negative. These data could support the presence of balancing selection at the β-globin gene cluster region in the Iranian population.

The results obtained from LD estimation indicated that D' values for three pairings of the markers, *Xmn*I-*Hind*IIIG, *Hind*IIIA-*Hind*IIIG and *Rsa*I-*Hinf*I, were higher than 0.5 in all three populations. These data indicate the presence of high linkage disequilibrium between these three paired markers. These LD values could provide useful information for selecting suitable markers in molecular diagnosis of beta thalassemia disease. The D' values close to zero for other pairings of markers showed that they were in linkage equilibrium. Therefore, the two-marker haplotypes *Xmn*I-*Rsa*I, *Hind*IIIA-*Rsa*I, *Hind*IIIG-*Rsa*I, *Xmn*I-*Hinf*I, *Hind*IIIG- *Hinf*I and *Hind*IIIA- *Hinf*I should be used with caution in diagnosis of disease.

Conclusion

In summary, the Iranian population showed heterozygosity deficiency for the RFLP markers including *Xmn*I, *Hind*IIIA, *Hind*IIIG, *Rsa*I and *Hinf*I, located within the β-globin gene region. Indeed, the haplotype phasing is more applicable than genotyping markers separately in the Iranian population. Although a large number of unrelated individuals from different provinces of Iran were included in the present study, one should note that the data cannot be applied to the whole Iranian population. More investigations on other parts of the Iranian population are required. The data obtained from this study could provide additional insight into the genetic structure of β-globin gene region in the Iranian population.

Acknowledgment

The authors declare that there is no conflict of interests.

References

1. Weatherall DJ, Clegg JB. Inherited haemoglobin disorders: An increasing global health problem. Bull World Health Organ 2001; 79:704-712.
2. da Luz JA, Sans M, Kimura EM, Albuquerque DM, de Fatima Sonati M, Costa FF. α-thalassemia, HbS, and β-globin gene cluster haplotypes in two Afro-Uruguayan sub-populations from northern and southern Uruguay. Genet Mol Biol 2006; 29:595-600.
3. Marengo-Rowe AJ. The thalassemias and related disorders. Proc (Bayl Univ Med Cent) 2007; 20:27-31.
4. Colosimo A, Guida V, De Luca A, Cappabianca MP, Bianco I, Palka G, *et al.* Reliability of dHPLC in mutational screening of β globin (HBB) genes. Hum Mut 2002; 19:287-295.
5. Goodman M, Koop BF, Czelusniak J, Weiss ML. The q globin gene: its long evolutionary history in the P-globin gene family of mammals. J Mol Biol 1984; 180: 803-823.
6. Easteal S. The pattern of mammalian evolution and the relative rate of molecular evolution. Genetics 1990; 124:165-173.
7. Gupta A, Sarwai S, Pathak N, Agarwal S. Beta-globin gene mutations in India and their linkage to β-haplotypes. Int J Hum Genet 2008; 8:237-241.
8. Derakhshan A, Karimi M, Ghadimi Moghaddam A. Comparative evaluation of renal findings in beta-thalassemia major and intermedia. Saudi J Kidney Dis Transpl 2008; 19:206–209.
9. Najmabadi H, Karimi-Nejad R, Sahebjam S, Pourfarzad F, Teimourian S, Sahebjam F, *et al.* The beta-thalassemia mutation spectrum in the Iranian population. Hemoglobin 2001; 25:285-296.
10. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988; 16:1215.
11. Sutton M, Bowhassira EE, Nagel RL. Polymerase chain reaction amplification applied to determination of β-like globin gene cluster haplotypes. Am J Hematol 1989; 32:66-69.
12. Guerreiro JF, Figueiredo MS, Santos SEB, Zago MA. β-globin gene cluster haplotypes in Yanomama Indians from the Amazon region of Brazil. Hum Genet 1992; 89:629-631.
13. Lancaster A, Nelson MP, Single RM, Meyer D, Thomson G. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. Pac Symp Biocomp 2003; 8:514-525.
14. Zhao JH. 2LD, GENECOUNTING and HAP: computer programs for linkage disequilibrium analysis. Bioinformatics 2004; 20:1325-1326.
15. Nassiry MR, Javanmard A, Tohidi R. Application of statistical procedures for analysis of genetic diversity in domestic animal populations. Am J Anim Vet Sci 2009; 4:136-141.
16. Wang J. Inbreeding and variance effective sizes for nonrandom mating populations. Evolution 1996; 50:1786–1794.