Molecular Basis of β-Thalassemia in Morocco: Possible Origins of the Molecular Heterogeneity

Imane Agouti,1 Catherine Badens,2 Ahmed Abouyoub,3 Nicolas Levy,2 and Mohcine Bennani1

We present the molecular spectrum of β-thalassemia in the Moroccan population obtained by the identification of molecular defects responsible for this disease, and herewith we show that the Moroccan population is genetically heterogeneous; 18 different mutations have been found in the 158 β-globin chromosomes studied. Eight mutations [codon 39 (C → T), FSC-8 (-AA), IVS-II-745 (C → G), −29 (A → G), FSC-6 (-A), IVS-I-110 (G → A), IVS-I-2 (T → C), and IVS-I-1 (G → A)] out of 18 β-thalassemia mutations identified accounted for 76% of the Moroccan β-thalassemia chromosomes. Restriction fragment length polymorphism (RFLP) haplotype analysis showed that the observed genetic diversity originated from both new mutational events and gene flow due to migration.

Introduction

β-Thalassemia is a group of heterogeneous autosomal recessive disorders due to absent or reduced synthesis of the β-globin chain (Weatherall and Clegg, 2001). The major pathophysiological determinant underlying β-thalassemia is the imbalance in the α/β-globin chain synthesis ratio; the excess of unmatched α-globins precipitates in the red cell precursors, damaging the membrane and leading to their premature destruction in the bone marrow. This ineffective erythropoiesis, combined with hemolysis in the periphery and an overall reduction in Hb synthesis (due to decreased β-chain production), leads to the severe anemia observed in β-thalassemia (Thein, 2005).

To date, more than 200 mutations, affecting different levels of β-globin gene expression, are known to result in a β-thalassemia phenotype (http://globin.bx.cse.psu.edu/hbvar) (Hardison et al., 2002; Patrinos et al., 2004). Since heterozygous β-thalassemia mutations provide a selective protection and advantage against malaria, β-thalassemia occurs with high frequencies among people living in regions where malaria is or has been endemic, such as Mediterranean countries, Africa, and Southeast Asia. In these populations, a limited number of alleles account for the majority of β-thalassemias, and only a small percentage of disease phenotype is due to a few rare mutations (Flint et al., 1998).

Being a North African country, Morocco is a melting pot of populations of Mediterranean, African, and even Asian origins. Various settlements have successively interspersed with Berbers throughout history: Phoenicians, Romans, and Arabs and, to a lesser degree, Greeks, Byzantines, and permanently Sub-Saharan Africans.

No precise epidemiological data are available for the Moroccan population; β-thalassemia carrier frequency is estimated to be between 1.5% and 3.0% (WHO, 1989). Only one molecular survey achieved in 2004 by Lemsaddek et al., combining results of Nadifi et al. (1996) and Lemsaddek et al. (2003), presented a spectrum of β-thalassemia mutations in the Moroccan population.

In the present report, a total of 158 β-globin chromosomes from patients selected from all over Morocco were studied. Our aim was to define the broad β-thalassemia mutational spectrum in this country, which is a prerequisite for defining a specific policy for carrier screening, genetic counseling, and, eventually, prenatal diagnosis. In addition, haplotype analysis was performed to indicate the possible origins of the detected β-thalassemia mutations.

Subjects and Methods

Subjects

The population sample consisted of 80 unrelated patients selected from different regions of Morocco. The diagnosis of β-thalassemia was based on clinical presentation of thalassemic features, further supported by relevant hematological data, as well as raised HbA2 levels in heterozygous family members. Blood samples were collected from patients during their attendance for blood transfusion in six major hospitals.
located in the different cities of the country, namely, Rabat, Casablanca, Tangier, Larache, Al-Hoceima, and Oujda. For haplotype analyses, DNA samples from the parents of the patients were also collected.

**DNA isolation and mutation screening**

Peripheral blood samples were collected in EDTA. DNA was extracted from leucocytes using the high salt method (Miller et al., 1988). Amplification refractory mutation system was used to screen for the presence of the common codon 39 (C→T) and IVS-I-1 (G→A) mutations (Old, 1996; Bravo et al., 1999).

The FSC-6 (-A) and the IVS-II-745 (C→G) mutations were characterized by PCR amplification followed by digestion with the restriction enzymes Bsu36I for the FSC-6 and Rsal for the IVS-II-745. The remaining mutations were characterized by direct sequencing of the amplified DNA, using an ABI PRISM® 3130 XL from Applied BioSystems (Foster City, CA).

**RFLP haplotype analysis**

β-Globin gene restriction fragment length polymorphism (RFLP) haplotypes were determined using PCR amplification of fragments containing seven polymorphic restriction sites located on the globin gene cluster. Following digestion with the specific restriction enzymes HincII 5’c, HincIII 5’Gγ, HindIII Aγ, HincII ψβ, HindII 3’ψβ, AzeII β, and HindI 3’β, samples were visualized by ethidium bromide staining after agarose or polyacrylamide gel electrophoresis and categorized according to Antonarakis et al. (1985).

**Results**

Of the total of 158 chromosomes investigated for β-globin gene mutations, 4 remained unknown, even after complete sequencing of their β-globin genes indicating probable deletion defects. The patients could be categorized as 76 homozygous β-thalassemia, 2 β-thalassemia carriers, and 4 β7/βthal cases.

Table 1 shows that the molecular basis of β-thalassemia in the Moroccan population is very heterogeneous with at least 18 different β-thalassemia alleles. Eight β-thalassemia mutations accounted for 76% of the Moroccan β-thalassemia chromosomes. These mutations included codon 39 (C→T), FSC-8 (AA), IVS-II-745 (C→G), −29 (A→G), FSC-6 (A), IVS-I-110 (G→A), IVS-I-2 (T→C), and IVS-I-1 (G→A). The remaining 10 alleles together were represented in 19% of the chromosomes at frequencies ranging between 3.16% and 0.63%. Thirteen of these were previously reported in Morocco (Lemsaddek et al., 2003), three are reported in Morocco for the first time [codon 24 (T→A), IVS-I-2 (T→C), and FSC-5 (G→A)], and two mutations have never been reported anywhere earlier. The novel mutations were a G→A substitution at position −190 relative to the β-globin gene cap site (Agouti et al., 2008) and the substitution A→G at position 726 of the second intervening sequence IVS-II (Agouti et al., 2007). Homozygosity for both rare and common mutations was very frequent in our cohort (~58%).

To determine the origin of mutant chromosomes in the Moroccan population, RFLP haplotype analysis of the 158 Moroccan chromosomes was performed. The association of β-globin RFLP haplotypes with the different mutations identified is presented in Table 1. The −29 (A→G) mutation is linked to four different haplotypes: II, IX, 3 black, and C. The mutations codon 39 (C→T) and FSC-8 (AA) are related to three haplotypes. Each of the IVS-I-110 (G→A), FSC-6 (A), IVS-I-2 (T→C), IVS-I-1 (G→A), and IVS-I-6 (C→T) mutations is linked to two haplotypes. The remaining 10 mutations were associated with a single haplotype (Table 1).
reflects the high rate of consanguinity of the Moroccan population described above. This considerable degree of homozygosity in spite of the high heterogeneity of mutations found in the Moroccan population will rise to 22, including new and rare mutations. Of the 80 subjects studied, 18 different \( -\)thalassemic patients reflects the ethnic diversity of the population. If these mutations are included, the number of \( -\)thalassemia alleles previously described in Morocco (Lemsaddek et al., 1997; Rosatelli et al., 1992b; 8, Waye et al., 1999; 9, Makhoul et al., 2005; 10, Tadmouri et al., 1998).

**Table 2. Frequency Distribution of the Most Common \( -\)-Thalassemia Mutations in Morocco and in Arab and Mediterranean Countries**

<table>
<thead>
<tr>
<th>Mutations</th>
<th>Morocco (160)</th>
<th>Algeria (239)</th>
<th>Tunisia (233)</th>
<th>Portugal (561)</th>
<th>Spain (324)</th>
<th>Italy (325)</th>
<th>Egypt (337)</th>
<th>Lebanon (520)</th>
<th>Turkey (795)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Codon 39 (C ( \rightarrow ) T)</td>
<td>26.58</td>
<td>27.6</td>
<td>40</td>
<td>36.8</td>
<td>36</td>
<td>40</td>
<td>1.5</td>
<td>0.5</td>
<td>3.8</td>
</tr>
<tr>
<td>FSC-8 (-AA)</td>
<td>13.91</td>
<td>–</td>
<td>0.9</td>
<td>–</td>
<td>0.4</td>
<td>0.1</td>
<td>1.8</td>
<td>2.5</td>
<td>5.4</td>
</tr>
<tr>
<td>IVS-II-745 (C ( \rightarrow ) G)</td>
<td>7.6</td>
<td>0.9</td>
<td>2.5</td>
<td>–</td>
<td>–</td>
<td>5</td>
<td>5.6</td>
<td>1.2</td>
<td>5</td>
</tr>
<tr>
<td>–29 (A ( \rightarrow ) G)</td>
<td>6.73</td>
<td>3.8</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FSC-6 (-A)</td>
<td>5.7</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>1.9</td>
<td>0.9</td>
<td>–</td>
<td>–</td>
<td>0.4</td>
</tr>
<tr>
<td>IVS-I-110 (G ( \rightarrow ) A)</td>
<td>5.7</td>
<td>24.7</td>
<td>20.5</td>
<td>10</td>
<td>13</td>
<td>19.9</td>
<td>32.9</td>
<td>34.2</td>
<td>39.2</td>
</tr>
<tr>
<td>IVS-I-2 (T ( \rightarrow ) C)</td>
<td>5.06</td>
<td>3.3</td>
<td>0.76</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>IVS-I-1 (G ( \rightarrow ) A)</td>
<td>5.06</td>
<td>11.7</td>
<td>1</td>
<td>28</td>
<td>325</td>
<td>10.2</td>
<td>11.3</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>89</td>
<td>72.3</td>
<td>75.8</td>
<td>85.4</td>
<td>77</td>
<td>54</td>
<td>53.4</td>
<td>58.8</td>
</tr>
<tr>
<td>References</td>
<td>a</td>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Values in parenthesis indicate the total number of chromosomes studied.

**Discussion**

In this paper, we describe the spectrum of \( -\)-thalassemia mutations in Morocco by analyzing 158 independent \( -\)-globin chromosomes. Along with the geographic distribution, haplotype analysis was conducted to study the origin of \( -\)-globin mutations. Eighteen different \( -\)-thalassemia mutations were identified, of which the IVS-II-726 (A \( \rightarrow \) G) and the –190 (G \( \rightarrow \) A) are novel mutations, and the codon 24 (T \( \rightarrow \) A), the IVS-I-2 (T \( \rightarrow \) G), and the FSC-5 (-CT) have never been reported before in Morocco. Conversely, four mutations previously described in Morocco (Lemsaddek et al., 2003) have not been identified in our subjects [IVS-I-130 (G \( \rightarrow \) A), the –101 (C \( \rightarrow \) T), the +20 (C \( \rightarrow \) T), and the poly A (T \( \rightarrow \) C)]. If these mutations are included, the number of \( -\)-thalassemia mutations found in the Moroccan population will rise to 22, reflecting even more the heterogeneous background of this population.

A significant feature of this study was the high prevalence of patients carrying the same mutation on both chromosomes, including new and rare mutations. Of the 80 subjects studied, more than a half (46) were true homozygotes (Table 1), in spite of the high heterogeneity of \( -\)-thalassemia alleles described above. This considerable degree of homozygosity reflects the high rate of consanguinity of the Moroccan population reaching 40% in certain regions (Talbi et al., 2006). The high consanguinity rate is mainly related to the high proportion of traditionally endogamous partner choice of the Moroccan population.

In contrast, the heterogeneity observed in the Moroccan thalassemic patients reflects the ethnic diversity of the population. Throughout history, Morocco has been a land of migration for several civilizations that contributed to its genetic admixture. The Berbers were the indigenous inhabitants; however, several waves of settlers of other ethnic groups have occurred. Around the 11th century BC, the Phoenicians settled in the land, extending their influence during a millennium. The Romans followed from the beginning of the Christian era until the 5th century AD when the Byzantine empires rolled until the conquest by Muslim Arabs during the 7th century AD. To these groups was added a continuous influx of Sub-Saharan Africans through the caravan routes. These successive waves of people throughout history have, no doubt, contributed to the genetic diversity of the current Moroccan population.

When compared to other countries (Table 2), the distribution of \( -\)-thalassemia mutations in Morocco was found to be generally similar; a small number of mutations predominated, and the most common ones were geographically the most widespread and theoretically the oldest. This was the case of the codon 39 (C \( \rightarrow \) T) mutation, which represented 26.58% of all Moroccan thalassemic chromosomes. This mutation of Western Mediterranean origin (Cao et al., 1989) reached its highest frequency in the ethnic isolation of Sardinia with 96% (Rosatelli et al., 1992a). It is the most frequent mutation in Algeria with 27.6% of \( -\)-thalassemic chromosomes (Bennani et al., 1994), in Tunisia and Italy with 40% (Fattoum et al., 1991; Rosatelli et al., 1992b; Haj Khelil et al., 2004), and in Spain with 36% (Amselem et al., 1988; Ribeiro et al., 1997). RFLP polymorphism analyses (Table 3) have demonstrated that the codon 39 mutation is associated with three haplotypes in Morocco and Algeria (Rouabbi et al., 1988), six in Tunisia (Haj Khelil et al., 2004), and nine different haplotypes in Sardinia (Piratsu et al., 1987). The association of the codon 39 mutation with several haplotypes can be explained not only by recombination events 5' to the \( -\)-globin gene but also by interallelic gene conversion in the case of haplotypes belonging to different \( -\)-globin frameworks. The heterogeneity of haplotype linkage in chromosomes carrying the codon 39 (C \( \rightarrow \) T) mutation argues for an ancient origin of this mutation in the Moroccan population, possibly Roman (Boletini et al., 1994). The next most common mutation (13.91% of alleles) was the 0-thal dinucleotide deletion (-AA) at codon 8. This mutation was originally detected in a Turkish patient (Orkin and Goff, 1981). The occurrence of this mutation in Ottomans was directly proven through the molecular analysis of the archeological remains of a child, who was homozygous for the FSC-8 mutation (Filon et al., 1995). This mutation is frequent in the Middle Eastern countries: 7.3% in Kuwait, 4.3% in the United Arab Emirates, and 5.4% in Turkey (El-Hazmi et al., 1995; Tadmouri et al., 1998; Zahed, 2001), but it is present in very low frequencies among Mediterraneans: 0.4% in Spain, 0.1% in Italy, and 0.9% in Tunisia (Fattoum et al., 1991; Rosatelli et al.,...
Table 3. Association of Mutations and Haplotypes in Morocco and in Arab and Mediterranean Countries

<table>
<thead>
<tr>
<th>Mutations</th>
<th>Morocco</th>
<th>Algeria</th>
<th>Tunisia</th>
<th>Portugal</th>
<th>Spain</th>
<th>Italy</th>
<th>Lebanon</th>
<th>Turkey</th>
</tr>
</thead>
<tbody>
<tr>
<td>This study</td>
<td>1,2*</td>
<td>3,4</td>
<td>5,6</td>
<td>7,8</td>
<td>9,10</td>
<td>11</td>
<td>12,13</td>
<td></td>
</tr>
<tr>
<td>FSC-8 (-AA)</td>
<td>IV, VI, VII</td>
<td>-</td>
<td>IV</td>
<td>-</td>
<td>-</td>
<td>IV, VII</td>
<td>IV, VII</td>
<td>IV, VII</td>
</tr>
<tr>
<td>FSC-7-II-745 (C → G)</td>
<td>VII</td>
<td>VII</td>
<td>-</td>
<td>VII</td>
<td>-</td>
<td>VII</td>
<td>VII</td>
<td></td>
</tr>
<tr>
<td>−29 (A → G)</td>
<td>II, IX, 3 black, C</td>
<td>IX, C</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>FSC-6 (-A)</td>
<td>III, IX</td>
<td>I, IX, A</td>
<td>IX, A, V</td>
<td>E</td>
<td>I, V</td>
<td>I, IX</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>IVS-I-110 (G → A)</td>
<td>I, II</td>
<td>I, II</td>
<td>I</td>
<td>I</td>
<td>I, II, IX</td>
<td>I, II, 5′–12</td>
<td>I, II, IV</td>
<td></td>
</tr>
<tr>
<td>IVS-I-2 (T → C)</td>
<td>IX</td>
<td>IX</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>IVS-I-1 (G → A)</td>
<td>V</td>
<td>I, III, V, IX, A</td>
<td>V</td>
<td>II, V, IX</td>
<td>II, V</td>
<td>V</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*1, Bennani et al. (1994); 2, Rouabhi et al. (1988); 3, Haj Khellil et al. (2004); 4, Chibani et al. (1988); 5, Faustino et al. (1999); 6, Cabeda et al. (1999); 7, Amselem et al. (1988); 8, Ribeiro et al. (1997); 9, Rosatelli et al. (1992b); 10, Firatsu et al. (1987); 11, Makhoul et al. (2005); 12, Diaz-Chico et al. (1988); 13, Aulehla-Scholz et al. (1990).

1992b; Benito et al., 1996; Haj Khellil et al., 2004). The finding that ~14% of chromosomes in our population carry this defect is unexpected because the Ottoman Empire never included Morocco. Our results show that the FSC-8 (-AA) mutation is associated with three different haplotypes, namely, IV [−−−−−−−], VI [−−−−−−+], and VII [−−−−−−−]. These haplotypes share the same Mediterranean framework 3 (Orkin et al., 1982); this type of mutation spread is most likely due to recombination events 5′ to the β-globin gene (Wong et al., 1986). Based on these results and the fact that one sequence haplotype was found to be associated with this mutation in the Turkish population (Tadmouri et al., 2001), a more extensive study of the genetic background of the FSC-8 (-AA) mutation, including sequence haplotypes, must be carried out in Morocco to explain the origin as well as the spread of this mutation. In addition to codon 39 and FSC-8, two other mutations were frequent in Morocco: IVS-II-745 (C → G) and −29 (A → G). The IVS-II-745 mutation was found to be predominant in Jordan (Sadiq et al., 2001) and in Syria (El-Hazmi et al., 1995), reaching a frequency of 14.2% and 16.7%, respectively; thus, this mutation could have been introduced into Morocco during the Arab conquests (8th–11th centuries) and then selected by malaria and endogamy. The −29 (A → G) mutation is believed to be of Sub-Saharan African origin and is specially frequent in Black Americans (Gonzalez-Redondo et al., 1991). Its presence in Morocco could be explained by migration during the Almoravid dynasty (1055–1130 AD) or through the caravan routes. The FSC-6 (-A) mutation reached its highest frequency in Algeria with 17% of the β-thalassemic chromosomes studied (Rouabhi et al., 1988; Bennani et al., 1994). It is almost always associated with haplotype IX, which is also more frequent in Morocco and Algeria (Rouabhi et al., 1988) than in other Mediterranean countries (Orkin et al., 1982). This would favor a North African origin of the FSC-6 mutation.

The IVS-I-110 (G → A) mutation was found at high frequencies not only in East-Mediterraneans, specially in Cyprus (77%) (Baysal et al., 1992) and Turkey (39.2%) (Tadmouri et al., 1998), but also in Algeria and Tunisia with 24.7% and 20.5%, respectively (Rouabhi et al., 1988; Fattoum et al., 1991; Bennani et al., 1994; Haj Khellil et al., 2004). The strong association of the IVS-I-110 mutation with the ancestral haplotype I suggested that it was possibly of ancient Greek origin (Cao et al., 1989). Our results show that IVS-I-1 (G → A) mutation is associated with haplotypes V and IX; however, it was linked to five RFLP haplotypes in Algeria (Rouabhi et al., 1988; Bennani et al., 1994). Based on the diversity of haplotypes associated with this mutation in Algeria, Bennani et al. (1994) and Perrin et al. (1998) have proposed a North African origin of IVS-I-1 mutation. Interestingly, IVS-I-1 was found to be predominant in Hungary (Ringelhann et al., 1993) and Czechoslovakia (Indrak et al., 1992), reaching a frequency of 29.4% and 45.2%, respectively; thus, it is possible that it may have originated in Eastern Europe. Nevertheless, in the absence of haplotype data for these populations, it is not possible to discuss a common or recurrent origin of the IVS-I-1 mutation. The IVS-I-2 (T → C) mutation is relatively frequent in Morocco with 5% of the β-thalassemic chromosomes studied. However, it is a rare allele observed sporadically in other parts of the world (Gonçalves et al., 1994). It was found at a high frequency (13%) in the Oranese population of Algeria (Bouhass et al., 1994) where it may have originated.

Conclusion

Considering the high mutational heterogeneity observed in Morocco, we believe that this study has provided useful information for both fundamental and practical purposes. The identification of the mutations occurring in this population represents a basis for the establishment of a β-thalassemia prevention program based on carrier screening and prenatal diagnosis. The genetic diversity observed indicates that both new mutational events and gene flow due to migration have occurred in the Moroccan population.

On the other hand, our data are obtained in a population with very high consanguineous marriage; therefore, the definition of independent chromosomes is bound to be an underestimation of the variability. Analysis of larger cohorts of patients not generated by consanguineous parents is bound to reveal even a larger spectrum of mutations than the present one.

Acknowledgments

This research was supported by PARS medecine N°178 and PROTARS II N° P12/29. We thank Dr. Danielle Lena-Russo, Centre d’Enseignement et de Recherche en Génétique Médicale, Faculté de Médecine Marseille, France, for her
cooperation during the course of these studies. The authors gratefully acknowledge Chantal Paolasso and Caroline Lacoste, Laboratoire de Génétique Moléculaire, Hôpital d’Enfant de la Timone, Marseille, France, for their excellent technical assistance. The authors also thank Pr. Khattab Mohamed, Service d’Hémato-Oncologie Pédriatique, Hôpital d’Enfant, Rabat, Maroc; Pr. Hachem and Dr. Assia Zakkarri Service de Pédiatrie III, Hôpital d’Enfant, Casablanca, Maroc; and Dr. Akodad Abid, Centre Régional de Transfusion Sanguine d’Al-Hoceima, for their help in obtaining blood samples.

Disclosure Statement

I declare, Agouti Imane, Doctor in Human Genetics, in the Faculty of Science and Technology of Tangier, Department of Biology, Laboratory of Applied Biology, on behalf of all co-authors of the manuscript named: “Molecular Basis of β-Thalassemia in Morocco: Possible Origins of the Molecular Heterogeneity” that no competing financial interests exist with any of the coauthors.

References


Address reprint requests to:

Agouti Imane, Ph.D.,
Laboratoire de Biologie Appliquée
Equipe de Génétique Moléculaire
Département de Biologie
Faculté des Sciences et Techniques
Bp 416 Tangier
Morocco

E-mail: imaneagouti@yahoo.fr