Abstract

β-thalassemia major (β –TM) is the most common thalassemia severe phenotype among Iranians. In recent years, molecular understanding of pathogenesis of β –TM has provided a great opportunity regarding diagnostic issues. Creating comprehensive molecular databases provides highly sensitive diagnostic tools for β –TM and effective prenatal diagnosis (PND) molecular screening tests. Despite a large body of research on molecular basis of β –TM, there are few review papers that consider a general view on the distribution of β –TM mutations in Iran. In the current review, common genetic defects identified in Iranian β –TM patients since 2005 to 2014 have been described. In addition, the prevalences and distributional trends of recognized mutations were discussed. It was found that IVSII-1 (G>A) and IVS I-5 (G>C) were by far the most frequent mutations detected in Iranian patients. Other common reported mutations included FSC 8/9 (+G), IVS I-110 (G>A), FSC 36/37 (–T), IVSI-1 (G>A), IVSI (-25bp), and codon 44 (-C). In conclusion, it was found that molecular profile of β –TM is highly variable among different Iranian populations; in particular, it seems that ethnicity and intra-migration can be most important participating factors in controlling distributional patterns.

Key words: β-thalassemia major, genetic modifiers, Iran, mutation

Introduction

β thalassemia major (β –TM) is the most common form of transfusion-dependent thalassemia in Iran. Most identified cases reside in south and north regions (more than 10 %), while other regions represent a mean frequency of 8-10 % (1, 2). Along with screening programs aimed to prevent thalassemia since 1991 in Iran, significant reduction was observed in emergence of new cases of β –TM in most regions of the country(3). However, we still face major problems controlling birth of new cases of β –TM in some regions (4). For instance, a high birth rate of new cases of β –TM has been reported in South-East province, Sistan and Baluchestan, mainly due to low educations (4). Including highly sensitive molecular tests in prenatal diagnosis (PND) screening programs provides an effective approach in order to manage birth of new β –TM cases. To apply in PND screening tests, however, molecular-based approaches need information on genetic background of specific populations in each region.

In the current review, a number of researches investigated β –TM mutations in Iranian populations from 2005 to 2014 were evaluated. Data from main geographical directions were gathered and entered into SPSS version 19 software. Statistical analysis was performed considering mean percentage of each mutation in individual areas. Finally, a brief discussion on latest identified molecular indicators of β-thalassemia intermediate phenotype was provided.

Beta globin gene cluster

In addition to adult beta globin gene (HBB), HBB cluster contains 4 other expressing genes namely epsilon, A-gamma, G-gamma, and delta. Each gene demonstrates unique expression
characteristics resulting in different hemoglobin profiles throughout life. Locus controlling region (LCR) roles as the main regulator sequence in switching between these active genes (Figure 1). For more details, reader is referred to comprehensive reviews on expression patterns of beta globin like genes (5). In addition to LCR, HBB contains a specific promoter sequence that provides higher levels of control on the gene expression. Furthermore, three exons and two introns (intervening sequence-IVS) are also present within the HBB which are targets for most mutations responsible for β−TM phenotype.

**Beta thalassemia mutations in Iran**

Majority of mutations resulting in β−TM are point mutations; however, deletions have also been reported. Up until, more than 300 mutations have been identified across the world; only a handful of them cover majority of observed mutations in specific geographical regions (7). Figure 2 demonstrates structure of HBB gene and commonly observed mutations. For more details on β−TM mutations, readers are referred to Thein 2013(6).

Owing to big geographical territory of Iran, multiple residing ethnicities and vast intra migrations, distribution of β−TM mutations varies considerably among Iranian populations (8-10). IVSII-1 (G>A) and IVSI-5 (G>C), two point mutations in intronic sequences of HBB gene, are by far the most frequently detected mutations in Iran (8,11,12). Other relatively common mutations include frame shift (FS) mutation in codons 8/9 (FSC 8/9 +G), IVSI-110 (G>A), FSC 36/37 (−T), IVSI-1 (G>A), IVSI (-25bp), and codon 44 (-C) (8). Less commonly observed mutations comprise codon 39 (C>T), codon 22 (−7bp), codon 5 (−CT), codon 30 (G>C), IVSI-II-745 (C>G), codon 8 (−AA), and IVSI-130 (G>C) (8, 11). Table I summarizes overall frequency of these mutations across the country, and table II shows relative frequency of the most common mutations in main geographical locations of the country.

Despite significant variety in molecular basis, mechanisms by which these defects blunt expression of HBB gene demonstrate more consistency. Majority of reported mutations within HBB gene interfere with mRNA processing (mainly mutations residing in intronic sequences). These mutations act either through altering constant di-nucleotides or consensus sequences in exon-intron junctions or activating of a non-effective cryptic mRNA splice site. Unlike most IVS mutations, majority of mutations detected within codon sequences create frame-shift alternations that interfere with mRNA translation (6). Nevertheless, responsible mutations cause a β0 (which has no mRNA or peptide output) phenotype in most cases (Table I).

**IVSII-1(G>A)**

Substitution of a single nucleotid (adenine instead of natural guanine base in the first nucleotide position of intron 2) causes IVSII-1 G>A mutation which is a null (β0) β-thalassemia allele (11). With a mean overall frequency of 23%, this mutation is the most frequent β-thalassemia mutation observed in Iran. However, its prevalence is not similar in all regions, with the highest frequency is observed in Northern regions (66%) (13). IVSII-1 G>A has also been reported at relatively high ratios in West, Central, North-West, and South-West of Iran. These regions are depicted in green color in Figure 3. Despite lower frequency of the mutation in South, South-East and North-East, IVSII-1 G>A still has been described among three most common mutations in these locations (Table II and Figure 4). Considering diverse distribution of IVSII-1 G>A in all regions of the country, including appropriate tests in PND screening programs able to identify this mutation is a necessity. A distributional trend is conceivable for IVSII-1(G>A) mutation. Because of
observing the highest frequency in north, it seems that this region is endemic site of the mutation followed by West, Central, North-West, and North-East areas (Figure 3). Regarding different resident ethnics in these regions, it seems that this similar distribution probably is result of inter migrations throughout the country.

**IVSI-5 (G>C)**

With Overall occurrence of 22%, IVS-I-5 (G>C) comprises the second most common mutation in Iranian patients. This mutation also results in a β0 thalassemia major phenotype, and interferes with normal mRNA splicing process. Regarding distributional pattern, the mutation is characteristic of South-East regions with a mean ratio of 73% (17, 20-22). IVS-I-5 (G>C) is also the most detected mutation in South of the country (66%) (22, 28). North-East comprises third common area of IVS-5(G>C) detection. Beside these three main geographical locations, IVI-5(G>C) mutation is also among three most common observed defects in patients of North and Central districts. However, this mutation is observed in lower ratios in Western lands (Table II).

Furthermore, IVSI-5 (G>C) is a common detected mutation in our neighbor countries. This mutation comprised the most common β-TM mutation in Pakistan (40.8%) (29). The mutation is also one of the frequent observed defects in some states of India (30, 31), Saudi Arabia, (32) United Arab emirate (33), Kuwait (34), Bahrain (34), and Iraq (35). In conclusion, distribution and frequency of this point mutation suggest the IVSI-5(G>C) mutation as a characteristic feature of Middle East β-TM population.

**FSC8/9, IVSI-110(C>T), FSC 36/37 (-T)**

Although far less common than IVSII-1(G>A) and IVSI-5(G>C), these mutations comprise third common mutational group (with average of 7.6-7.8%) among Iranian β-TM patients. FSC 8/9 +G (FS resulting from a guanine base insertion in joining position of codons 8 and 9) is a β0 allele, and accounts for the most common reported FS mutation in these patients. Nevertheless, FSC 8/9 +G along with IVSI-110 (C>T), which also affects mRNA splicing process and creates a severe form of β+ phenotype, are highest detected mutations in North-West region (24% and 21% respectively, Table II). Following by North-West are Western regions (with overall mean ratio of 12% and 10% for FSC 8/9+G and IVSI-110 (C>T)) respectively (10, 24, 26, 27). Other areas have been reported with lower penetrances for these mutations. A mean ratio of 3-8% has been observed in Northern, Eastern, Southern, and Central regions. Lowest rates have been described in South with 0.3% frequency for FSC 8/9+G and South-east with 1% occurrence for IVSI-110 (C>T). Table II shows frequencies of these mutations in different geographical locations. Another relatively frequent FS mutation, FSC 36/37 (-T), has been identified with highest rate in West (20.8%) and South-West (14%). Following by these regions are central areas with a mean ratio of 7% penetrance of FSC 36/37 (-T). Frequency of the recent mutation is as low as 0.5% in other regions (Table II).

**IVSI-1(G>A), IVSI (--25bp) and FSC44 (-C)**

These mutations are relatively common in Iranian patients. IVSI-1(G>A) creates a β0 allele, and similar to the most other mutations residing in IVS sequences, this mutation also suppresses normal mRNA processing. Showing mean overall frequency of 4.3%, IVSI-1(G>A) represents as a relatively common encountered β-TM mutation in Iran. Highest frequency of this mutation was reported from North-West (13.2%), while lowest was described in South-East (0.8%). IVSI-1(G>A) has also been detected in West and South-West regions with 5% penetrance. However, the mutation is rarely seen in patients of North and South origin.
Deletion of a 25 bp sequence at 3' end of IVSI (IVSI (-25bp)) causes a β₀ thalassemia allele interfering with normal mRNA splicing. This mutation has been seen with a mean ratio of 4.8% among Iranian β-TM cases. However, maximum rate of the mutation (12.3%) has been described in South-West, while minimum rate (0.3%-0.7%) was noted in North-west, North-East, and South-East regions.

FSC 44(-C) mutation is resulted from a single base deletion (C) at codon 44 of HBB gene, and creates a β₀ allele. This mutation comprises a mean prevalence of 3.9% across the country; however, relatively high entrance (24%) of the mutation has been observed in North-East (Table II). Lowest rates of FSC 44(-C) have been related to South-East and Western regions (range of 0.3-0.9%).

**Other common mutations**

Base substitution (G>T) in Codon 22 (β₀ allele), an adenine base deletion (-A) within Codon 6 (β₀ allele), C>T substitution in Codon 39 (β₀ allele), C>G substitution in IVSII-745 position (β⁺ allele), FS resulting from C base deletion at Codon 16 (β₀ allele), and a double base (-CT) deletion at Codon 5 (β₀ allele) with overall frequencies ranging from 1-4 % are less commonly encountered mutations in Iranian β-TM patients. Codon 22 (G>T) presents as a relatively common mutation in North, however, overall penetrance of the mutation is low across the country. This mutation has only been reported from South-West regions with a low percentage (0.5%). Codon 6 (-A) with a mean overall frequency of 2.7% has been reported from approximately all geographical areas. Highest rate of Codon 6 (-A) has been revealed in West (5.9%) with an overall mean frequency of 2.1%. Another less observed mutation in Iranian patients is codon 16 (-C). This defect occurs with general ratio of 2.5%, and shows similar distribution in different areas of the country. In contrast to most mRNA processing mutations in Iranian patients, IVSII-745 (C>G) is a relatively rare mutation. With a net mean frequency of 2.3%, IVSII-745 (C>G) along with Codon 39 (C>T) are responsible for 4.6% of β-TM mutations in Iran. These two mutations also show similar distributions among different country zones. Double base deletion (-CT) at codon 5 resulting in a non-functional FS in codon reading pattern is a rare mutation among our patients. Other commonly detectable mutations include FSC 8 (–AA) and FSC 15(G>A). These defects are among common mutations reported from West (10). Likewise, FSC 8(–AA) is among frequent mutations found in North-West (8%) (16). Nonsense point mutation in codon 15(G>A) producing a β₀ allele covers 5 % of responsible mutations in South-East. In addition, a base change (C>T) in codon -88 which creates a mild β⁺⁺ allele was reported as a relatively common mutation in South-East (3.4 % frequency).

**Molecular determinants of Phenotype and outcome**

Polymorphisms in specific genes are involved in improvement of clinical course of β-TM. Boosting Hb F level is regarded as the primary strategy of ameliorating mechanisms; however, some studies have reported alleviation of clinical severity of β-TM without significant enhancement in Hb F level (36). These observations highlight possible role of unknown molecular mechanisms participating in controlling clinical severity of the disease. Xmn-I gamma (γ) globin gene polymorphism (presence of G nucleotide rather than A nucleotide at 5’HS4-LCR palindromic polymorphic site), co-inheritance of α thalassemia and inheritance of β⁺ alleles account for majority of intermediate phenotypes in β-TM. In addition, δ-β deletion and co-inheritance of HbS have also been described to create an intermediate phenotype (37, 38). In spite of availability of sensitive molecular assays, no genetic
causes are identifiable in around 20% of β-thalassemia intermediate patients (38).

**Gamma globin gene alternations**

As mentioned, Xmn-I γ globin gene is regarded as a strong phenotype modifier in β-TM patients (37-40). This polymorphism has been observed to be associated with higher Hb F production and lower disease complications (40, 41). In addition, an association was found between Xmn-I polymorphism and appropriate response to hydroxyl urea treatment (42). Nevertheless, there are conflicting results regarding association of Xmn-I polymorphism and response to hydroxyl urea (43, 44). These contradict results highlight possible role of other unidentified mediator mechanisms beside Hb F over-expression which may be responsible for amelioration of β-TM phenotype (43, 44). An association of Xmn-I and BCL11A gene polymorphisms was suggested as a genetic profile reducing phenotype intensity in β-TM. In fact, BCL11A polymorphism has displayed no impacts on clinical severity of patients with A allele at 5'HS4-LCR of γ globin gene (45).

In addition, variations in -588 position of Aγ globin allele have been supposed to participate in Hb F silencing processes in adults. This alternation has been demonstrated to be associated with Xmn-I polymorphism in thalassemia intermediate patients (39). Current understandings of molecular mechanisms involved in HbF silencing are depicted in Figure 5.

HS-11(-21 A), another polymorphism in promoter region of γ globin gene, has been proposed as a genetic marker of β-thalassemia intermediate (39). As well, Lou JW et al identified an association between γ globin gene triplication and a 1.3 kb length base deletion in β globin gene cluster to improve phenotype of β-TM patients (46).

**Other modifications in beta globin locus**

Some mutations in β globin gene are characterized with mild reduction in output of mRNA synthesis, and therefore result in a mild clinical picture of the disease. IVS-I-6 (T > C), Codon -88 (C > A), and Codon +113 (A > G) comprise common mild alleles reported in relation to thalassemia intermediate phenotype in Iranian patients (37, 38). Moreover, compound heterozygous inheritance of Hb Knossos (HBB:c.82G>T; Codon 27 GCC→TCC (Ala→Ser)) with β0 mutation; IVSII-1 (G>A), has been reported in association with a thalassemia intermediate phenotype (47). Interestingly, a β0 FS mutation, (HBB: c.44delT (p.Leu14ArgfsX5) was recently reported to act as a genetic modifier of TM phenotype (48).

Coinheritance of polymorphisms within the erythroid Krüppel-like factor (EKLF) with Hb E (HBB: c.79G>A) and α+-thalassemia was described to cause β-thalassemia intermediate phenotype (49). Recently, mutations in KLF1 were found in 12 β-thalassemia intermediate patients who had significant transfusion-free survival periods (50). These observations suggest the potential role of genetic alternations in erythroid specific transcription factors as possible genetic modifiers in β-TM patients. Currently, a novel 11kbp deletion within β-LCR regions that removes a sequence between DNase I hypersensitive site 2 (HS2) and HS4 of the β-LCR has been described to results in a εγδβ0-form of thalassemia presenting with normal hematological and clinical phenotype (51).
Table I: Characteristics of common observed β-thalassemia mutations in Iranian patients in different locations along with pathogenesis of each mutation

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Mean frequency</th>
<th>Type of allele (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVS-II-1(G&gt;A)</td>
<td>23</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>IVS-I-5 (G&gt;C)</td>
<td>22</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codons 8/9 (+G)</td>
<td>7.8</td>
<td>β°, RNA translation</td>
</tr>
<tr>
<td>IVS-I-110 (G&gt;A)</td>
<td>7.6</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codon 36/37(-T)</td>
<td>7.6</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>IVSII-25 bp Deletion</td>
<td>4.8</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>IVSI-1 (G&gt;A)</td>
<td>4.3</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codon 44(-C)</td>
<td>3.9</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codon 22(- 7bp)</td>
<td>3.9</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codon 6(-A)</td>
<td>2.7</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codon 16(-C)</td>
<td>2.5</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>IVSII-745(C&gt;G)</td>
<td>2.3</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codon 39 (C&gt;T)</td>
<td>2.3</td>
<td>β°, RNA processing</td>
</tr>
<tr>
<td>Codon 5(-CT)</td>
<td>1.3</td>
<td>β°, RNA processing</td>
</tr>
</tbody>
</table>

Table II: Observed percentages of common Iranian β-thalassemia mutations in 8 main geographical locations. Empty cells denote no available data

<table>
<thead>
<tr>
<th></th>
<th>IVSII-1(G&gt;A)</th>
<th>IVSI-5 (G&gt;C)</th>
<th>FSC8/9 (+G)</th>
<th>IVSI-110 (G&gt;A)</th>
<th>FSC36/37 (-T)</th>
<th>IVSI (-25bp)</th>
<th>IVSI-1 (G&gt;A)</th>
<th>FSC44 (-C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>66</td>
<td>13.6</td>
<td>3.6</td>
<td>4.8</td>
<td>0.5</td>
<td>2.2</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>North-East</td>
<td>8</td>
<td>42.9</td>
<td>4.7</td>
<td>3.6</td>
<td>0.7</td>
<td>24.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South-East</td>
<td>4</td>
<td>73.1</td>
<td>4.2</td>
<td>1</td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>South</td>
<td>9</td>
<td>66.8</td>
<td>0.3</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South-West</td>
<td>16</td>
<td>5.3</td>
<td>5.8</td>
<td>9.6</td>
<td>14.8</td>
<td>12.3</td>
<td>5.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Central</td>
<td>27</td>
<td>9.1</td>
<td>6.1</td>
<td>9.6</td>
<td>7.4</td>
<td>5.1</td>
<td>2.4</td>
<td>3.4</td>
</tr>
<tr>
<td>West</td>
<td>30</td>
<td>3.4</td>
<td>12.3</td>
<td>10</td>
<td>20.8</td>
<td>5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>North-West</td>
<td>25</td>
<td>4.8</td>
<td>24.3</td>
<td>21</td>
<td>0.3</td>
<td>13.2</td>
<td>3.2</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. β globin gene cluster on chromosome 11. Locus control region (LCR) which contains four DNA hypersensitivity (HS) sites is the major site for interaction of erythroid specific transcription factors. Four active genes including ε, Gγ, Aγ, δ and β are mapped. Each of these genes expresses in certain periods of human development with the β gene as the main expressing gene in adulthood. Also, a pseudo β -like gene is located between Aγ and δ genes.(adapted from Thein et al, 2013)(6).

Figure 2. Map of human β-globin (HBB) gene and common indicated mutations along with their location, and their effects on gene production. Majority of β-thalassemia mutations in world are point mutations resulting in abnormal processing of mRNA or early blockage of transcription process. Tow most common mutations in Iranians occur in nucleotide 1 of IVS-I and nucleotide 5 of IVSII. Figure is adopted from Thein; 2013.
Figure 3. Distributional patterns of the two most common $\beta$-thalassemia mutations in Iranian population; IVSII-1(G>A) (yellow) and IVSI-5(G>C) (blue). IVSII-1(G>A) mutation which is a null $\beta^0$ allele demonstrates overall frequency of 23% across the country. IVSII-1(G>A) occurs with highest prevalency in North region, following by North-west, West, Central and in lower rate in South-East regions. Unlike IVSII-1(G>A), IVSI-5(G>C) mutation produces a $\beta^0$ allele with highest prevalence seen in South-East. Lower ratios of this mutation have been reported from South and North-East. A lower percentage of IVSI-5(G>C) mutation was observed in North region which mostly is result of migration of southern residents in the past years. Specific ratios of the both mutations are represented in table 2. These specific patterns highlight the role of inter migration in mutational profiles rather than ethnicity effect.
Figure 4. Three most common β-thalassemia mutations observed in each region of Iran. IVSII-1(G>A) comprises the most common detected mutation with overall rate of 23% in Iranian patients. This mutation has had the first mutational rank in 5 main geographical directions including North, West, Central, North-West and South-West, in order to frequency. Also, this mutation comprises the second and the third common mutation in other 3 geographical locations, namely South, South-East and North-East. IVSII-1(G>A), as the second most common mutation, was reported with highest rate in South-East, South and North-East. Unlike IVSII-1(G>A), IVSII-1(G>C) has been only observed as the second and the third common allele in North and Central regions respectively. Other than these two splice site mutations, a frame-shift interfering with mRNA translation has comprised the third common β-thalassemia mutation in Iran. Overall, these three main mutations comprised the following ratios of all observed mutations: 84.4% in North, 75.7% in North-East, 81.3% in South-East, 79.1% in South, 43.1% in South-West, 45.7% in Central, 63.1% in West and 70.3% in North-West.
Figure 5 Molecular targets identified as effectors in silencing of gamma globin gene expression in adults. BCL11A is the main mediator that its polymorphisms are known modifiers of β-thalassemia major phenotype. Other main contributing transcription factors include Kruppel like factor 1, c-MYC, and GATA binding factor 1. However, exact role of demonstrated factors on γ gene expression is still unclear. Figure is adopted from Thein; 2013.

Conclusion
Although responsible mutations for β-TM are heterogenous among Iranian β-TM patients, limited numbers of mutations are responsible for majority of genetic defects in different areas of the country. This fact provides the possibility of planning efficient PND platforms for diagnosis of β-TM using molecular based approaches in each region.

Newly identified molecular determinants affecting clinical course of β-thalassemia are progressively becoming clear. Although role of Xmn-I polymorphism of γ globin gene has been established as a strong genetic determinant, researches highlight role of other mechanisms involving in this process.

Conflict of interest
The authors declare no conflict of interests in the production of this manuscript.

References
24. Mehrabi M, Alibakhshi R, Fathollahi S, Farshchi MR. The Spectrum of β-Thalassemia Mutations in


