Bioinformatics

Beta Thalassemia

Sami Khuri
Department of Computer Science
San José State University
San José, California, USA
sami.khuri@sjsu.edu
www.cs.sjsu.edu/faculty/khuri

Beta globin gene

Uppercase characters:
- mature mRNA

Lowercase characters:
- introns
- flanking sequences

Red
- cat box
- tata box
- polyadenylation sequence

Table 1 Types and Frequencies of β-Thalassemia Mutations and β-Globin variants in Lebanon

Mutation	Phenotype	Number of Chromosomes	Number of homozygotes/heterozygotes	Frequency (%)
IVS-I-110 (G>A)	β^+	178	65/48	34.2
IVS-I-1 (G>A)	β^0	78	29/20	15.0
IVS-I-6 (T>C)	β^+	75	28/19	14.4
cd 29 (C>T)	β^+	50	22/6	9.6
IVS-II-1 (G>A)	β^0	45	17/11	8.6
cd 5 (−CT)	β^0	26	9/8	5.0
cd 30 (G>C)	β^0	14	6/2	2.7
cd 8 (−AA)	β^0	13	5/3	2.5
cd 44 (−C)	β^0	8	3/2	1.5
IVS-II-745 (C>G)	β^+	6	3/0	1.1
β^0	β^0	5	0/5	1.0
−87 (C>G)	β^+	4	1/2	0.8
IVS-I-5 (G>C)	β^+	4	2/0	0.8
−88 (C>T)	β^+	3	1/1	0.6
290 bp deletion	β^0	3	1/1	0.6
25 bp deletion	β^0	2	1/0	0.4
$\delta\beta$-thalassemia (Sicilian type)	$\delta\beta^0$	2	1/0	0.4
cd 8/9 (+G)	β^0	1	0/1	0.2
cd 36/37 (−T)	β^0	1	0/1	0.2
cd 39 (C>T)	β^0	1	0/1	0.2
Unknown		1	0/1	0.2
Total		520	194/132	100

"Genetic heterogeneity of beta thalassemia in Lebanon reflects historic and recent population migration" by N. J. Makhoul, et al.
"Annals of Human Genetics" in 2005 (issue 69, pages 55 to 66).

Table 1 Types and Frequencies of β-Thalassemia Mutations and β-Globin variants in Lebanon

Mutation	Phenotype	Number of Chromosomes	Number of homozygotes/heterozygotes	Frequency (%)
IVS-I-110 (G>A)	β^+	178	65/48	34.2
IVS-I-1 (G>A)	β^0	78	29/20	15.0
IVS-I-6 (T>C)	β^+	75	28/19	14.4
cd 29 (C>T)	β^+	50	22/6	9.6
IVS-II-1 (G>A)	β^0	45	17/11	8.6
cd 5 (−CT)	β^0	26	9/8	5.0
cd 30 (G>C)	β^0	14	6/2	2.7
cd 8 (−AA)	β^0	13	5/3	2.5
cd 44 (−C)	β^0	8	3/2	1.5
IVS-II-745 (C>G)	β^+	6	3/0	1.1
β^0	β^0	5	0/5	1.0
−87 (C>G)	β^+	4	1/2	0.8
IVS-I-5 (G>C)	β^+	4	2/0	0.8
−88 (C>T)	β^+	3	1/1	0.6
290 bp deletion	β^0	3	1/1	0.6
25 bp deletion	β^0	2	1/0	0.4
$\delta\beta$-thalassemia (Sicilian type)	$\delta\beta^0$	2	1/0	0.4
cd 8/9 (+G)	β^0	1	0/1	0.2
cd 36/37 (−T)	β^0	1	0/1	0.2
cd 39 (C>T)	β^0	1	0/1	0.2
Unknown		1	0/1	0.2
Total		520	194/132	100

Six mutations are responsible for 86.8% of all beta thalassemias in Lebanon

"**Genetic heterogeneity of beta thalassemia in Lebanon reflects historic and recent population migration**" by N. J. Makhoul, et al.
"Annals of Human Genetics" in 2005 (issue 69, pages 55 to 66).

TABLE 2. FREQUENCY DISTRIBUTION OF THE MOST COMMON β-THALASSEMIA MUTATIONS IN MOROCCO AND IN ARAB AND MEDITERRANEAN COUNTRIES

Mutations	*Morocco (160)*[a]	*Algeria (239)*	*Tunisia (233)*	*Portugal (561)*	*Spain (324)*	*Italy (325)*	*Egypt (337)*	*Lebanon (520)*	*Turkey (795)*
Codon 39 (C→T)	26.58	27.6	40	36.8	36	40	1.5	0.5	3.8
FSC-8 (-AA)	13.91	–	0.9	–	0.4	0.1	1.8	2.5	5.4
IVS-II-745 (C→G)	7.6	0.9	2.5	–	–	5	5.6	1.2	5
−29 (A→G)	6.33	3.8	–	–	–	–	–	–	–
FSC-6 (-A)	5.7	17	6.65	1	1	1.9	0.9	–	0.4
IVS-I-110 (G→A)	5.7	24.7	20.5	10	13	19.9	32.9	34.2	39.2
IVS-I-2 (T→C)	5.06	3.3	0.76	–	–	–	–	–	–
IVS-I-1 (G→A)	5.06	11.7	1	28	35	10.2	11.3	15	5
Total	76	89	72.3	75.8	85.4	77	54	53.4	58.8
References[b]		1	2,3	4	5,6	7	8	9	10

[a]Values in parenthesis indicate the total number of chromosomes studied.
[b]1, Bennani *et al.* (1994); 2, Fattoum *et al.* (1991); 3, Haj Khelil *et al.* (2004); 4, Faustino *et al.* (1999); 5, Amselem *et al.* (1988); 6, Ribeiro *et al.* (1997); 7, Rosatelli *et al.* (1992b); 8, Waye *et al.* (1999); 9, Makhoul *et al.* (2005); 10, Tadmouri *et al.* (1998).

"**Molecular basis of beta thalassemia in Morocco: possible origins of the molecular heterogeneity**" by I. Agouti et al. In "Genetic Testing"
Volume 12, Number 4, 2008.

TABLE 2 Spectrum of Relative Frequencies (%) of β-Thalassemia Mutations in Tunisia, Algeria and Morocco

Mutation	Type	Origin	Tunisia			Algeria		Morocco	
References			44	29	33	47	46	42	48
Codon 39, C>T	β^0	Mediterranean	48.76	43.78	49.00	25.90	27.60	26.58	26.20
IVS-I-110, G>A	β^+	Mediterranean	15.70	10.81	21.00	26.40	24.70	4.70	3.21
Codon 8, –AA	β^0	East European	–	0.54	0.20	0.90	–	13.91	9.63
Codon 6, –A	β^0	North African	1.65	7.02	2.60	12.90	17.10	5.70	13.37
IVS-I-6, T>C	β^+	West Mediterranean	2.48	0.54	0.60	6.20	3.30	3.16	13.90
IVS-I-1, G>A	β^0	Asian	3.31	4.32	4.50	9.10	11.70	5.06	8.56
Codon 5, –CT	β^0	Greek	–	–	0.40	–	–	1.27	–
Codon 44, –C	β^0	Kurdish Jews	1.65	1.62	3.80	–	–	–	–
IVS-I-2, T>C	β^0	Algerian; Russian	–	0.54	–	0.90	3.30	5.06	2.14
IVS-I-2, T>G	β^0	Tunisian	19.83	0.54	3.00	–	–	3.16	0.53
–30, T>A	β^+	Turkish	–	–	0.80	–	0.40	–	–
–29, A>G	β^+	Black	–	1.08	–	1.40	3.80	6.33	4.28

"**Hemoglobinopathies in North Africa**" by A. Haj Khelil et al. In "Hemoglobin, 34" Vol 34, Numb 1, 2010.

Major Intermedia and Minor

- The alpha and beta loci determine the structure of the 2 types of polypeptide chains in adult hemoglobin, Hb A.
- Absence of beta chain causes beta-zero-thalassemia.
- Reduced amounts of detectable beta globin causes beta-plus-thalassemia.
- For clinical purposes, beta-thalassemia is divided into:
  - thalassemia major (transfusion dependent),
  - thalassemia intermedia (of intermediate severity), and
  - thalassemia minor (asymptomatic).


```
ATG GTG CAT CTG ACT CCT GAG GAG AAG TCT GCC GTT
ACT GCC CTG TGG GGC AAG GTG AAC GTG GAT GAA GTT
GGT GGT GAG GCC CTG GGC AG  GTTGGTATCAAGGTTACAA
GACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGA
AGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTC
TATTTTCCCACCCTTAG G CTG CTG GTG GTC TAC CCT TGG
ACC CAG AGG TTC TTT GAG TCC TTT GGG GAT CTG TCC
ACT CCT GAT GCT GTT ATG GGC AAC CCT AAG GTG AAG
GCT CAT GGC AAG AAA GTG CTC GGT GCC TTT AGT GAT
GGC CTG GCT CAC CTG GAC AAC CTC AAG GGC ACC TTT
GCC ACA CTG AGT GAG CTG CAC TGT GAC AAG CTG CAC
GTG GAT CCT GAG AAC TTC AGG GTGAGTCTATGGGACGCTT
GATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAGG
AAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAAT
```


IVS-1-6 (T > C)

IVS-1-1 (G > A)

IVS-1-110 (G>A)

```
ATG GTG CAT CTG ACT CCT GAG GAG AAG TCT GCC GTT
ACT GCC CTG TGG GGC AAG GTG AAC GTG GAT GAA GTT
GGT GGT GAG GCC CTG GGC AG GTTGGTATCAAGGTTACAA
GACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGA
AGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTC
TATTTTCCCACCCTTAG G CTG CTG GTG GTC TAC CCT TGG
ACC CAG AGG TTC TTT GAG TCC TTT GGG GAT CTG TCC
ACT CCT GAT GCT GTT ATG GGC AAC CCT AAG GTG AAG
GCT CAT GGC AAG AAA GTG CTC GGT GCC TTT AGT GAT
GGC CTG GCT CAC CTG GAC AAC CTC AAG GGC ACC TTT
GCC ACA CTG AGT GAG CTG CAC TGT GAC AAG CTG CAC
GTG GAT CCT GAG AAC TTC AGG GTGAGTCTATGGGACGCTT
GATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAGG
AAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAAT
```


cd 5 (-CT)

cd 29 (C>T)

IVS-II-1 (G > A)

```
ATG GTG CAT CTG ACT CCT GAG GAG AAG TCT GCC GTT
ACT GCC CTG TGG GGC AAG GTG AAC GTG GAT GAA GTT
GGT GGT GAG GCC CTG GGC AG GTTGGTATCAAGGTTACAA
GACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGA
AGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTC
TATTTTCCCACCCTTAG G CTG CTG GTG GTC TAC CCT TGG
ACC CAG AGG TTC TTT GAG TCC TTT GGG GAT CTG TCC
ACT CCT GAT GCT GTT ATG GGC AAC CCT AAG GTG AAG
GCT CAT GGC AAG AAA GTG CTC GGT GCC TTT AGT GAT
GGC CTG GCT CAC CTG GAC AAC CTC AAG GGC ACC TTT
GCC ACA CTG AGT GAG CTG CAC TGT GAC AAG CTG CAC
GTG GAT CCT GAG AAC TTC AGG GTGAGTCTATGGGACGCTT
GATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAGG
AAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAAT
```

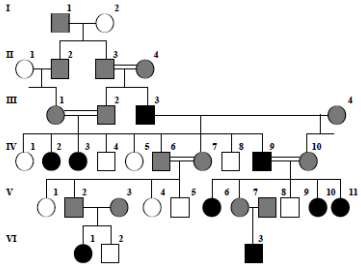

YYYYYYNYAG|G

IVS-1-110 (G>A)

```
ATG GTG CAT CTG ACT CCT GAG GAG AAG TCT GCC GTT
ACT GCC CTG TGG GGC AAG GTG AAC GTG GAT GAA GTT
GGT GGT GAG GCC CTG GGC AG GTTGGTATCAAGGTTACAA
GACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGA
AGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTC
T ATT TTC CAC CCT TAG G CTG CTG GTG GTC TAC CCT TGG
ACC CAG AGG TTC TTT GAG TCC TTT GGG GAT CTG TCC
ACT CCT GAT GCT GTT ATG GGC
```

Consensus: YYYYYYNYAG|G
Mutation: TGCCTATTAG|T
Wild Type: TGCCTATTGG|T

Consanguineous Family



In the **Western family**:
* Mutation carriers are usually scattered throughout the general population.
So the genetic basis for their disease predisposition is missed
In the **Arab family**:
* Arab society mutation carriers mostly remain concentrated within the extended family,
so the genetic nature of their disease predisposition may be obvious.

Genetic Disorders in the Arab World: United Arab Emirates
The Arab World by Ghazi Omar Tadmouri

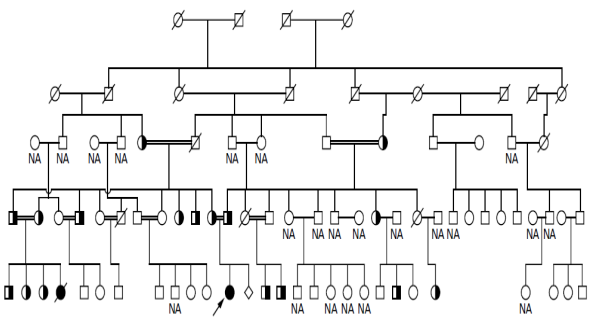
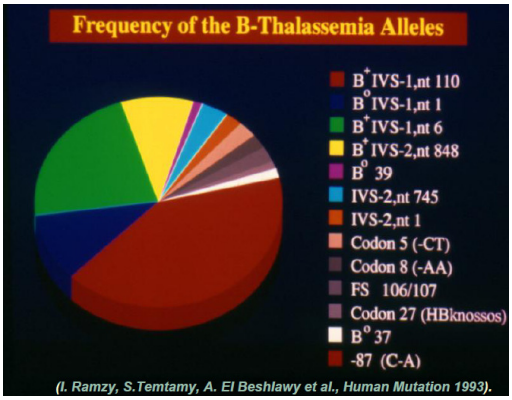


Figure 1. Pedigree and the Results of Testing for β-Thalassemia in a Typical Index Family.

Screening extended families for genetic hemoglobin disorders in Pakistan, by S. Ahmed et al., Vol. 347, No. 15, October 2002, The New England Journal of Medicine



Prevention of Thalassemia

Problems with the prevention of thalassemia in Egypt

- National screening for thalassemia not yet available.
- Centers for genetic counseling and premarital examination are lacking.
- Prenatal diagnosis centers are few and not yet well equipped.
- Prenatal diagnosis conducted on a limited and voluntary basis.
- Population awareness is improving.

Free Of Thalassemia by 2012

- The first project of the UAE GDA (Emirates Free Of Thalassemia by the year 2012) is aimed at identifying the Beta-Thalassemia and Sickle-cell carriers in the UAE pre-marital population.
- UAE GDA seeks to step up the fight against Thalassemia, in line with the UAE federal government's vision to make the country free from the new births of children with Thalassemia major by the year 2012.

Molecular Defects in Thalassemia

The molecular defects identified in thalassemias:

- gene deletion, e.g., of the terminal portion of the beta gene
- chain termination (nonsense) mutations
- point mutation in an intervening sequence
- point mutation at an intervening sequence splice junction
- frameshift deletion
- fusion genes, e.g., the hemoglobins Lepore; and
- single amino acid mutation leading to very unstable globin,
  - Example: Hb Vicksburg (beta 75 leu-to-0).