Sequence Analysis of Human HBA and HBB Thalassemia Genes

Shahid. R[1], Muhammad. W. S[2], Sadaf. S[3] , Saima. J[4] and Hira. M[5]

Abstract— Thalassemias are diverse group of hereditary disorders in which there is a low rate of synthesis of one or more of the globin polypeptide chains. Also, one of most common autosomal recessive disorders worldwide. Thalassemia are quantitative abnormalities of polypeptides globin chain synthesis. Normally an individual inherits two β-globin genes located one each on two chromosomes 11, and two alpha globin genes one each on two chromosome 16, from each parent i.e. normal adult hemoglobin is α2β2. The study is focused on analysis of mutated genes, their sequences and protein domains. The human alpha (HBA1) and beta (HBB) loci determine the structure of the 2 types of polypeptide chains in adult hemoglobin HbA.

Index Terms— Hereditary disorder, Autosomal, Polypeptides, Domain.

1 INTRODUCTION

Hemoglobin disorders constitute the most common lethal inherited disorders worldwide. Thalassemia is the commonest inherited hemoglobinopathy[1]. 'Thalassaemia' are genetic disorders and have a widespread distribution in many countries of the world. Thalassaemias were among the first diseases to be characterized at the molecular level and inherited in a mendelian recessive fashion. The severe, homozygous form of the disease is called thalassaemia major, and the carrier state, in which only one defective globin gene is inherited is called the trait. Most of the thalassaemias result from mutations that involve either the alpha or beta globin genes. Alpha thalassemia occurs when a gene or genes related to the alpha globin protein are missing or changed. Beta thalassemia occurs when similar gene defects affect production of the beta globin protein.

1.1 Molecular Pathogenesis

Each hemoglobin molecule contains four subunit proteins. Two of the subunit proteins are called alpha and two are called beta. Hemoglobin properly binds and releases oxygen only when two alpha subunits are connected to two beta subunits [2,3] A pair of genes located on chromosome #16 controls the production of the alpha subunits of hemoglobin. A single gene located on chromosome #11 controls the production of the hemoglobin beta subunit. Thalassemia occurs when one or more of the genes fail to produce protein, leading to a shortage of one of the subunits.

1.2 Thalassemia Major and Minor

Individuals with thalassemia major usually present within the first two years of life with severe anemia, requiring regular red blood cell (RBC) transfusions[4,5]. Findings in untreated individuals with thalassemia major are growth retardation, pallor, jaundice, poor musculature and skeletal changes that result from expansion of the bone marrow. Carriers of beta-thalassemia are clinically asymptomatic[6]. The characteristic hematological features are microcytosis (reduced red blood cell volume), hypochromia (reduced red blood cell Hb content) and increased HbA2 level.

[1]Dean, Faculty of Biotechnology and Agro- Sciences, University of South Asia, Lahore, Pakistan. .*PH-92-3004238303. E-mail: mianrs@yahoo.com*

[2]Resident/Researcher, Hera General Hospital, Makkah, Kingdom of Saudi Arabia/ National institute for Biotechnology and genetic engineering, Faisalabad, Pakistan.

PH-92-3006630086. E-mail: mwshoaib@hotmail.com

[3]Resident, District Head Quarter (DHQ), Faisalabad, Pakistan. *PH-92-3216671049*

[4]Resident/DMS, District Head Quarter (DHQ), Faisalabad, Pakistan. *PH-92-3219663538. E-mail:saima_sh4e@hotmail.com*

[5]Lecturer, Faculty of Biological Sciences, University of South Asia, Lahore, Pakistan*PH-92-3007943269. E-mail: hira_sh@hotmail.com*

1.3 Beta Thalassemia

Beta-thalassemia is one of most common autosomal recessive disorders worldwide. Different molecular mechanisms like base substitutions, small deletions or insertions of one or two nucleotides in the β-globin gene are responsible for β-thalassemia [7]. Moreover, it has been found that β-thalassemia mutations are relatively population specific [8-11]. β-thalassemia is reported as the single-gene disorder.

2 Material and Methods

The alpha (HBA1, 141800 HBA2, 141850) and beta (HBB) loci determine the structure of the 2 types of polypeptide chains in adult hemoglobin, HbA. Mutant beta globin that sickles causes sickle cell anemia. Absence of beta chain causes beta-zero-thalassemia. Reduced amounts of detectable beta globin causes beta-plus-thalassemia.

2.1 Mutation analysis

The β-thalassemias can be caused by more than 200 different HBB gene mutations. The prevalent molecular defects are limited in each at-risk population. The most commonly used methods are reverse dot blot analysis or primer specific amplification ARMS PCR, real-time PCR or microarray technology.

2.2 Bioinformatics Approach For Sequence Analysis

Sequence analysis detects mutations in the HBB coding region and associated flanking regions. Different bioinformatics softwares were used to analyze the selected sequences.

2.3 Transcript Analysis

Transcript maps were obtained through Ensemble software.http://www.ensembl.org/Homo_sapiens/Transcript/Sequence_cDNA?db=core;g=ENSG00000206172;r=16:176680-177522;t=ENST00000320868.

3 Results

The reported symbol for hemoglobin beta gene was HBB with HGNC ID: 4827. The HBB gene was located on chromosome 11. Table 1 Shows the HBB gene Symbol with other details.

Table 1

Symbol Report: HBB

APPROVED SYMBOL	HBB
APPROVED NAME	hemoglobin, beta
HGNC ID	HGNC:4827
PREVIOUS SYMBOLS & NAMES	-
SYNONYMS	beta-globin, CD113t-C, HBD
LOCUS TYPE	gene with protein product
CHROMOSOMAL LOCATION	11p15.5
GENE FAMILY	Hemoglobin subunits
HCOP	Orthology Predictions for HBB

HBB Sequence Retrieval

DEFINITION Homo sapiens hemoglobin, beta (HBB), mRNA
ACCESSION NM_000518, mRNA 626 bp

GCTGCCTTGGATACCGCCCTCTAGGTATCGTIRGCGIGTTTTTATBTATACCGGC-
TACATGCBAAGGCCCCTAAGTAGOCGGGCTSAAAICAPTCAAGGCAAC-
GTAAGGATCCCG-
GOTTTTCTAACTCG8GALGTGCGCCTTCAAAAACGCCCACGCGGTC1ATTCGGCACTGC
CGTTTCCCCGTGAMTGTAGTTNTGCTGATGATGCCTT-
GTNG0GT0CATCAGATGGAGTG2CG6GTCATBGTCSCTTATAAGCGCTACGGGIATGAAT
CCTGGCAGGGGTACCCTCMTCGAGGCTAGCTGTGFTGATAAGTCAACTCGTCAAATT-
GGC_AATCTCC6CGGGGAATTAAAGTGTNTAGCAAA0PCTGHAACGTCA,AANOCGCAA
GCGGTCCATCACBGAGCTAGCGCCTGDCSATATGGAA-
GANTGCGTTCAAATCCATGTTCGCATTAGGCCATCAA-
TA5AGATTGACTCACACACTACTCGTGRTAICNGTGGTATTATOMATTGTTGGTCTTTG
TTAATCAAGCGGEGTGGGTTT-
GTAGETTGGCGCTBTAGGCCA,,GAGCCGGEAITCCCCTAGAGTCATTTGCCGGATGCTC
GGATATACGOGGGTANTOTGTTACCCNGGACGCCCCAACGACGTEGMTTGMTTAC-
GCTGCSHGATCCATTCTCGTAGACTGTTCACHTCCAGECTCA

The highlighted area shows the coding region (Exon))

HBB transcript Map analysis via Ensemble

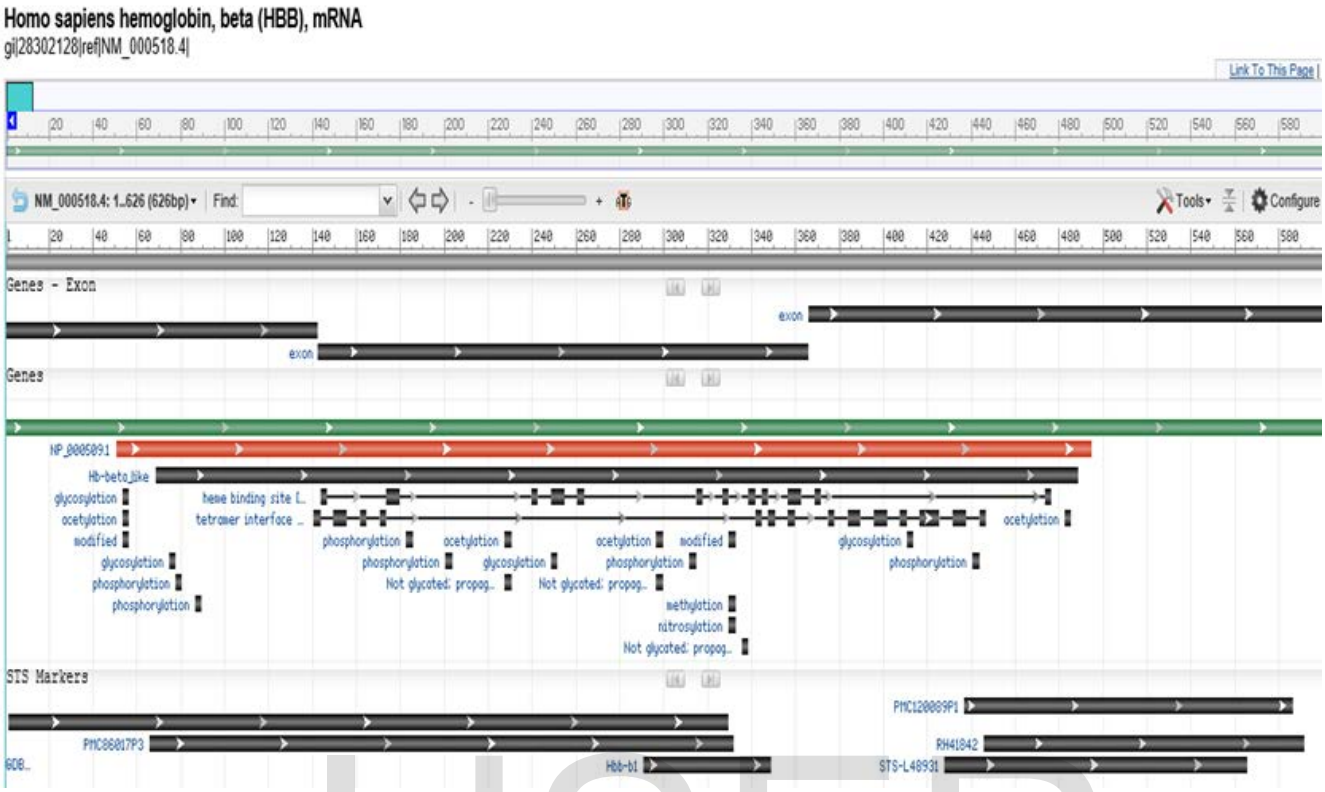
Gene: HBB ENSG00000244734

Description	hemoglobin, beta [Source:HGNC Symbol;Acc:HGNC:4827]
Synonyms	beta-globin, CD113t-C, HBD
Location	Chromosome 11: 5,225,464-5,229,395 reverse strand. GRCh38:CM000673.2
About this gene	This gene has 5 transcripts (splice variants), 136 orthologues, 9 paralogues, is a member of 2 Ensembl protein families and is associated with 25 phenotypes.
Transcripts	Hide transcript table

Show/hide columns (1 hidden) Filter

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq	Flags
HBB-001	ENST00000335295	628	147aa	Protein coding	CCDS7753	D9YZU5 P68871	NM_000518 NP_000509	TSL:1 GENCODE basic APPRIS P1
HBB-004	ENST00000380315	502	90aa	Protein coding	-	F8W6P5	-	CDS 3' incomplete TSL:5
HBB-005	ENST00000633227	609	55aa	Nonsense mediated decay	-	-	-	
HBB-002	ENST00000485743	680	No protein	Retained intron	-	-	-	TSL:1
HBB-003	ENST00000475226	319	No protein	Retained intron	-	-	-	TSL:2

Protein domain analysis of HBB gene



HBA Gene

The reported symbol for hemoglobin beta gene was HBA with HGNC 4823. The HBA gene was located on chromosome 16. Table 2 Shows the HBB gene Symbol with other details.

Table 2

Symbol Report: HBA1

APPROVED SYMBOL	HBA1
APPROVED NAME	hemoglobin, alpha 1
HGNC ID	HGNC:4823
PREVIOUS SYMBOLS & NAMES	-
SYNONYMS	HBA-T3
LOCUS TYPE	gene with protein product
CHROMOSOMAL LOCATION	16p13.3
GENE FAMILY	Hemoglobin subunits
HCOP	Orthology Predictions for HBA1

HBA Sequence Retrieval

DEFINITION Homo sapiens hemoglobin, alpha 1 (HBA1), mRNA.
ACCESSION NM_000558, mRNA 627 bp

GGGCCTACCCGTACCCTGAGCTTGCGAGTCCACCGGGGCACCCATCAGTACGCCGACTT-
GCGGTACACCCCGAACAACGCGTAGCAAGCTTAAGTGTCTGACGGTAAGTCGTTTGCCTGAG-
TTGGGGGCGCACCCCCCCCCCTGCGCCCTTATTCTTCC-
TATCCGATTGAGCTGGCCCGCCGGCCAGAGTCTCAGTTCCCAAGGTGCTCCTCTCACGGG-
CATGGCTCACCAGAGGACACCAGCAACCAGGACCCGATTAATGCAC-
CTATCCTCCCCGCGCGTTCCGCGCCAGTAGGGTCGGAATTCACTCAAAGCCTTTAGCGAC-
CTGGTCGCCTCCTGCCCTCACACGGGACGCAATCAGATCACCGACCTGACGCCGCTCGG-
CAGATGTCAAGCTATCACCACGCGCCTACGCCAACCCCCTATTGTGCGCGCGAG-
TCCGCTCCCTTCGGCCGCGCGGGTATCCAGCATAACGGTGTTCGGGGAACCATCACAGGAG-
CACCTTCAATGAGTCACGATCTCCGCTCGAGAGGGCTGCGGCGCCAAACGGGCCAC-
CGGCGCTCGCCGTACCCCGAACCTCCTAGCAAGGGGCCGCTG-
TAAACTGTCGCCAACAAAAGGAGACGACTTGC

(The highlighted area shows the coding region (Exon))

HBB transcript Map analysis via Ensemble

Gene: HBA1 ENSG00000206172

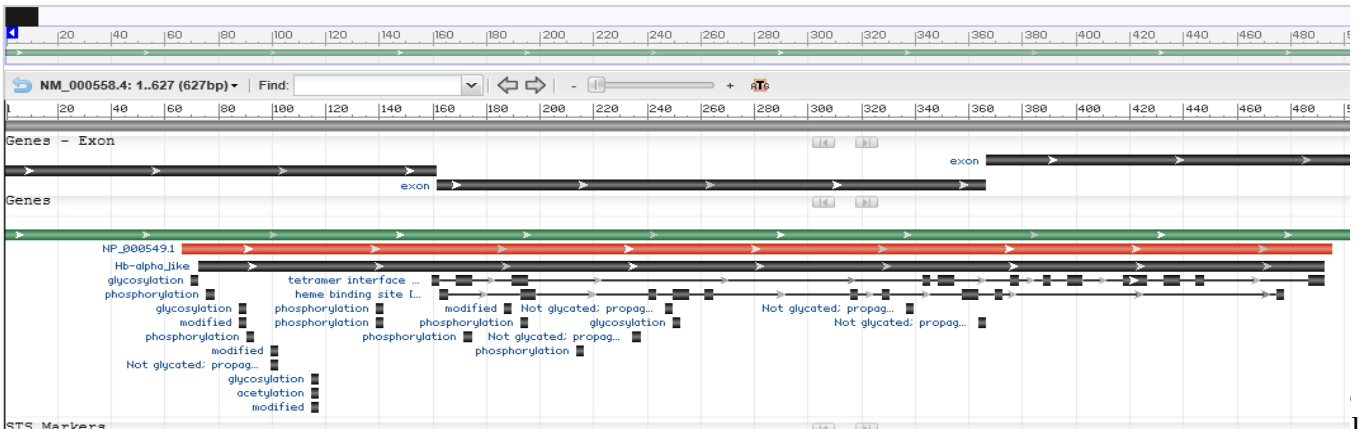
Description	hemoglobin, alpha 1 [Source:HGNC Symbol;Acc:HGNC:4823]
Synonyms	HBA-T3
Location	Chromosome 16: 176,680-177,522 forward strand. GRCh38:CM000678.2
About this gene	This gene has 4 transcripts (splice variants), 54 orthologues, 9 paralogues, is a member of 1 Ensembl protein family and is associated with 4 phenotypes.
Transcripts	Hide transcript table

Show/hide columns (1 hidden) Filter

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq	Flags
HBA1-001	ENST00000320868	577	142aa	Protein coding	CCDS10399	D1MGQ2 P69905	NM_000558 NP_000549	TSL:1 GENCODE basic APPRIS P1
HBA1-006	ENST00000397797	504	110aa	Protein coding	-	G3V1N2	-	TSL:2 GENCODE basic
HBA1-002	ENST00000472694	674	No protein	Retained intron	-	-	-	TSL:1
HBA1-005	ENST00000487791	410	No protein	Retained intron	-	-	-	TSL:2

Protein domain analysis of HBB gene

Homo sapiens hemoglobin, alpha 1 (HBA1), mRNA
gi|672228742|ref|NM_000558.4|



sub types. Over the past three decades, regular blood transfusions and iron chelation have dramatically improved the quality of life and transformed thalassemia from a rapidly fatal disease in early childhood to a chron-

ic disease compatible with prolonged life [12,13]. Today life expectancy varies between 25-55 years, depending on the compliance with medical treatment [54,55]. The fact that the number of patients with thalassemia decreases beyond 15 years could be explained by death mostly among children older than 15 years. This can be explained by the fact that if children are not transfused, they die before the age of 6 years and if they are transfused and non-chelated, they die before the age of 20.

5 CONCLUSION

Mutations causing the thalassemias have now been characterized and a small number of common mutations cause the bulk of disease in each particular population-base. With little more than 6 to 8 common mutations probes over 90% of thalassemic patients can now be characterized. But challenges remain in the 10% where the mutations are rare, or have not yet been determined. Newer developments in micro array technology in combination with current PCR based systems will lead to further characterisation and aid proper genetic potential identification and control of the disorder. Gene therapy is an exciting prospect. Deliveries of transgenes in stem cell based gene therapy are effective in the therapeutic management. Sequence analysis through Bioinformatics approaches shows coding regions of HBB and HBA genes. Protein sequence analysis shows various domians in each sequence. To study, sequence variations and to analyse these at molecular level can help to overcome many genetic disorders.

ACKNOWLEDGMENT

We are thankfull to all colleagues for their support and fruitfull suggestions.

REFERENCES

[1] Weatherall DJ, Clegg JB. Oxford: Blackwell Scientific Publications; 2001. The Thalassaemia Syndromes.
[2] Vichinsky, E; Levine, L and Bhatia, S et al. (2008), “Standards of Care Guidelines for Thalassemia”, Children’s Hospital & Research Center Oakland, U.S.A.
[3] SL, Schrier (2002), “Pathophysiology of Thalassemia”, Curr. Opin. Hematol., 2,123126.
[4] Old J, Traeger-Synodinos J, Galanello R et al. Prevention of thalassemias and other hemoglobin disorders. Volume 2: Laboratory Methods. Published by Thalassemia International Federation (TIF) 2005.
[5] 5. Ye BC, Zhang Z, Lei Z. Molecular analysis of alpha/beta-thalassemia in a southern Chinese population. Genet Test 2007; 11: 75–83. 18.
[6] Vrettou C, Traeger-Synodinos J, Tzetis M et al. Rapid screening of multiple beta-globin gene mutations by real-time PCR on the Light Cycler: application to carrier screening and prenatal diagnosis of thalassemia syndromes. Clin Chem 2003; 49: 769–76.
[7] Huisman THJ, Carver MFH, Baysal E. A Syllabus of Thalassemia Mutations. The Sickle Cell Anemia Foundation, Augusta, GA 1997.
[8] Furuumi H, Firdous N, Inoue T et al. Molecular basis of beta thalassemia in the Maldives. Hemoglobin 1998; 22: 141-51.
[9] Setianingsih I, Williamson R, Dand D et al. Phenotypic variability of Filipino beta(0)-thalassemia/HbE patients in Indonesia. Am J Hematol 1999; 62: 7-12.
[10] Colah R, Gorakshakar A, Nadkarni A et al. Regional heterogeneity of beta-thalassemia mutations in the multi ethnic Indian population. Blood Cells Molecules Dis 2009; 42: 241-6.
[11] Colah R, Gorashekar A, Phanasgaonkar S et al. Epidemiology of beta thalassemia in western India: mapping the frequencies & mutations in subversion of Maharashtra and Gujrat. Br J Haemato 2010; 149: 739-47.
[12] Hahalis G, Alexopoulos D, Kremastinos DT et al. Heart failure in beta thalassemia syndromes: a decade of progress. Am J Med 2005; 118: 957-67.

[13] Pearson HA, Cohen AR, Giardina PJ et al. The changing profile of homozygous beta thalassemia: demography, ethnicity, and age distribution of current North American patients and changes in two decades. Pediatrics 1996; 97: 352-6.

[14] Satwani H, Raza J, Alam M et al. Endocrinal complications in thalassemias: Frequency and association with ferritin levels. Pak Ped J 2005; 29: 113-19.

[15] Agarwal MB. Advances in management of thalassemia (Editorial). Indian Pediatrics 2004; 41: 989-92.