

## Research Article

# In silico mutation analysis of human beta globin gene in sickle cell disease patients

Hira Mubeen<sup>1,3\*</sup>, Rubab Zahra Naqvi<sup>3</sup>, Ammara Masood<sup>3</sup>, Muhammad Waseem Shoaib<sup>2</sup>,  
Shahid Raza<sup>1</sup>

<sup>1</sup>University of South Asia, Lahore, Pakistan

<sup>2</sup>District Head Quarter (DHQ), Faisalabad, Pakistan

<sup>3</sup>National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan

**Received:** 10 March 2016

**Revised:** 24 April 2016

**Accepted:** 07 April 2016

### \*Correspondence:

Dr. Hira Mubeen,

E-mail: [hira\\_sh@hotmail.com](mailto:hira_sh@hotmail.com)

**Copyright:** © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ABSTRACT

**Background:** Sickle cell disease is an inherited blood disorder that affects red blood cells. People with sickle cell conditions make a different form of hemoglobin called hemoglobin S. Sickle cell conditions are inherited from parents in much the same way as blood type, hair color and texture, eye color and other physical traits. Sickle cell disease occurs due to a single mutation on the b-globin gene, namely, a substitution of glutamic acid for valine at position 6 of the b chain. Several mutations in HBB gene can cause sickle cell disease. Abnormal versions of beta-globin can distort red blood cells into a sickle shape. The sickle-shaped red blood cells die prematurely, which can lead to anemia. The study is focused on analysis of HBB gene with its different variants, Evolutionary pathways and protein domains by using various bioinformatics tools.

**Methods:** The study is focused on analysis of HBB gene with its different variants, Evolutionary pathways and protein domains by using various bioinformatics tools.

**Results:** Sickle cell disease occurs due to a single mutation on the b-globin gene, namely, a substitution of glutamic acid for valine at position 6 of the b chain. Several mutations in HBB gene can cause sickle cell disease. Abnormal versions of beta-globin can distort red blood cells into a sickle shape. Comparative study shown 38 different genes with little genetic variation among different species.

**Conclusion:** Studies suggested that there is need to maintain a primary prevention program to detect sickle cell disease at earlier stages despite having a large high risk. Preventive diagnosis and follow-up would reduce infant mortality by preventing the development of severe anemia as well as dangerous complications. In short, sickle cell disease surveillance would avert loss of life, measured as the number of years lost due to ill-health, disability or early death.

**Keywords:** Substitution, Sickle shaped, Hemoglobin, Evolutionary pathway

## INTRODUCTION

Sickle Cell Anemia is a severe illness that affects millions of people all across the globe. Approximately 2 million Americans carry the sickle cell trait. The overall

incidence of SCD is eight out of 100,000 people. However, it is much more widespread in some people. The genetic defect that causes sickle cell anemia affects hemoglobin. Hemoglobin is a constituent of red blood cells that carries oxygen to all the cells and tissues in the

body. 'Red blood cells that contain normal hemoglobin are soft and round. People with SCD, however, have a type of irregular hemoglobin. 'A genetic error makes the hemoglobin molecules stick together in a long, rigid rods after they release oxygen. These rods cause the red blood cells to become hard and sickle-shaped, unable to squeeze through tiny blood vessels. Various studies suggested that 250,000 children are born annually with sickle cell anemia worldwide and thus it is among the most important epidemiological genetic diseases in Brazil and the world.<sup>1-2</sup> The disease occurs due to a mutation of the beta globin gene of hemoglobin, causing a substitution of the glutamic amino acid for valine at position 6 of the beta chain, resulting in production of an abnormal hemoglobin, called hemoglobin S (Hb S), instead of normal hemoglobin, hemoglobin A (Hb A). With modified physicochemical characteristics, the molecules of hemoglobin S suffer polymerization and precipitation, leading to a change in form, a deformity of red blood cells which become sickle-shaped.<sup>3</sup> The inheritance of sickle cell anemia occurs via an autosomal recessive gene with both parents, working as an asymptomatic carriers of a single affected gene (heterozygous), transmitting the defective gene to their homozygous child (Hb SS).<sup>4</sup> In the beginning, it was reported by some scientist that the sickle gene spreads by migration of a single mutation.<sup>5</sup> Later on, results of restriction fragment length polymorphism analysis on the beta-globin gene cluster indicates the sickle gene mutation may have developed independently and spontaneously at least five times.<sup>6-7</sup> Although the molecular abnormality leading to the sickle gene is the same in all haplotypes. Hence, a wide variation in the clinical manifestations and severity of the associated disease was observed. Due to the fact, the clinical phenotype of SCD is said to be multigenic.<sup>8</sup>

### **Genetic Modification**

Some previous studies showed the diverse effects with use of some vectors. It was observed that the development of some integrating vectors for  $\beta$ -globin gene transfer has been challenging due to the complex regulatory elements needed for high-level, erythroid-specific expression.<sup>9</sup>  $\gamma$ -Retroviral vectors were unable to transfer these  $\beta$ -globin expression cassettes intact.<sup>10-11</sup> Whereas the lentiviral vectors (LV) can transfer  $\beta$ -globin cassettes intact with relatively high efficiency.<sup>12-13</sup> In the last decade, many groups have developed different  $\beta$ -globin LV for targeting  $\beta$ -hemoglobinopathies, with successful therapeutic results following transplantation of ex vivo-modified HSC in mouse models.<sup>14-15</sup> Another useful approach is to modify  $\beta$ -globin genes to confer anti-sickling activity by substituting key amino acids from  $\gamma$ -globin. The modified  $\beta$ -globin cassette should yield the necessary high-level, erythroid-specific expression in adult erythroid cells. An LV carrying a human  $\beta$ -globin gene with the amino acid modification T87Q was designed by Pawliuk et al.<sup>16</sup> The glutamine at position 87 of  $\gamma$ -globin has been implicated in the anti-

sickling activity of HbF.<sup>17</sup> This anti-sickling construct corrected SCD in 2 murine models of the disease, and a similar LV has been used in a clinical trial for  $\beta$ -thalassemia and SCD in France.<sup>18</sup>

### **Prevalence**

According to World Health Organization report, the most valid measure to study the impact of SCD on public health is under-5 years old mortality. SCD contributes the equivalent of 5% of under-5 deaths in the African continent, more than 9% of such deaths in West Africa, and up to 16% of under-5 deaths in individual West African countries. An increasing number of affected children currently survive five years of age but remain at risk of premature death, and 48% of patients surviving into adulthood have chronic organ dysfunction.

## **METHODS**

### **Bioinformatics approach for Sequence Analysis**

Sequence analysis detects mutations in the HBB coding region and associated flanking regions. The HBB gene provides instructions for making beta-globin. Different mutations are caused due to various forms of beta globin in the HBB gene. One particular HBB gene mutation produces an abnormal version of beta-globin known as hemoglobin S (HbS). Different bioinformatics software's were used to analyze and compare the selected Beta globin gene sequence. Gene sequences were analyzed using NCBI web server.

### **Phylogenetic Analysis**

Phylogenetic analysis was done using CLUSTALW software. [http://www.ebi.ac.uk/Tools/services/web\\_clustalw2\\_phylogeny/toolform.ebi](http://www.ebi.ac.uk/Tools/services/web_clustalw2_phylogeny/toolform.ebi)

### **Protein domain analysis**

Protein domain for beta globin gene was analyzed using CDD tool from NCBI.

### **Pfam structure analysis**

Protein domains were analyzed using EMBL-EBI Pfam database. <http://pfam.xfam.org/family/Globin>.

## **RESULTS**

Sequence analysis detects various genetic variations including polymorphisms in the coding region of HBB gene and associated flanking regions. Different bioinformatics software's were used to analyze the selected HBB gene sequence. Comparative study shown 38 different genes with little genetic variation among different species.

Sequence Analysis

Sequence analysis detects various genetic variations including polymorphisms in the coding region of HBB gene and associated flanking regions. Results are shown below for HBB gene with comparison in other species.

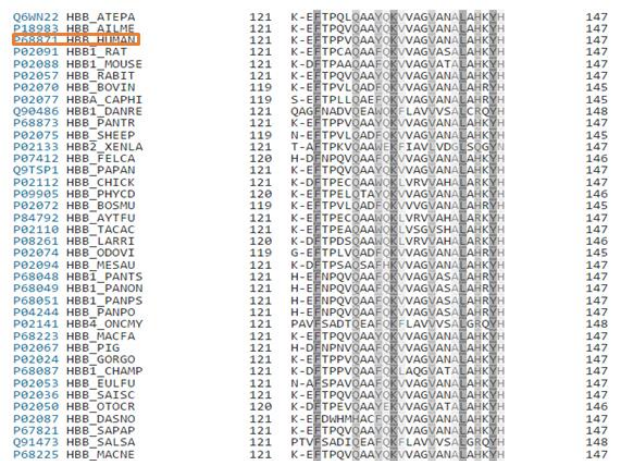
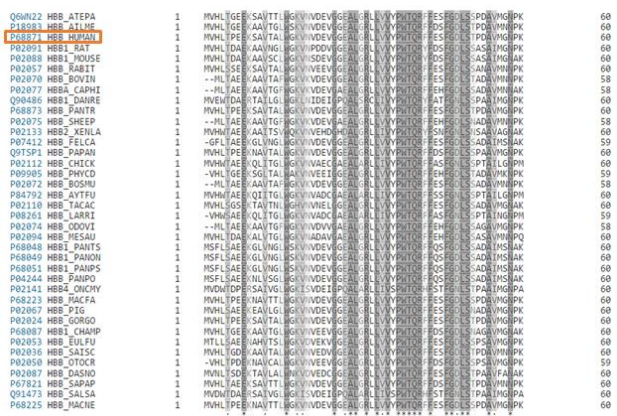
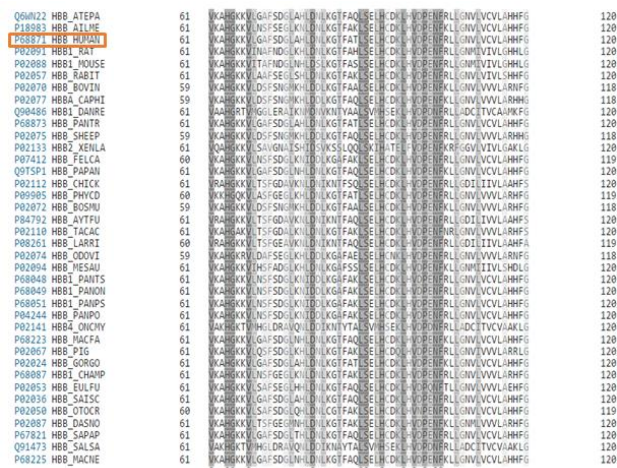


Figure 1: Sequence alignments of beta globin gene across different species using clustalw.

Phylogenetic Analysis

Phylogenetic analysis is used to estimate the evolutionary relationships. The evolutionary history inferred from phylogenetic analysis is usually depicted as branching, treelike diagrams that represent an estimated pedigree of the inherited relationships among species. Below is the diagrammatic tree representation for globin gene with other species.

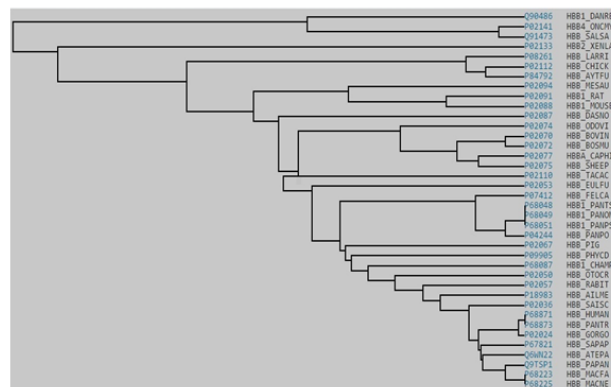


Figure 2: Phylogenetic tree of beta globin gene across different species using clustalw.

pfam structure of globin domain

The globins are a family of globular proteins which are thought to share a common ancestor. These proteins all incorporate the globin fold, a series of eight alpha helical segments. Two prominent members of this family include myoglobin and hemoglobin, which both bind the heme prosthetic group. Both of these proteins are reversible oxygen binders. Structural domain was obtained using pfam database. Below is the visual representation of globin domain.

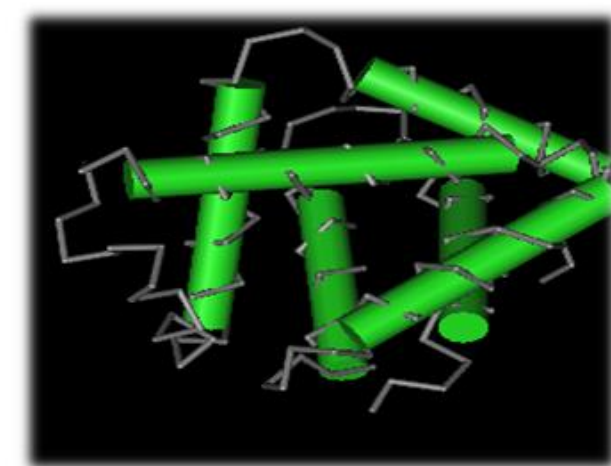
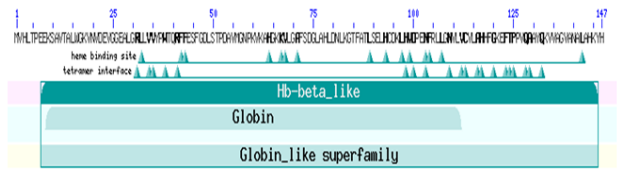


Figure 3: pfam structure of globin domain.

### Domain pf beta globin protein

Protein domains were analyzed using conserved domain software. Results shown that the heme binding site starts at position 32 and ends at 142. Similarly, the tetramer interface starts at position 31 and ends at 132. Below is the clear representation of HBB globin superfamily domain.



**Figure 4: CDD predicted domain PF beta globin protein.**

### DISCUSSION

Sickle cell disease occurs due to a single mutation on the b-globin gene due to a substitution of glutamic acid for valine at position 6 of the b chain. Persons with sickle cell trait (SCT) are heterozygous carriers of an abnormal  $\beta$ -globin gene that results in the production of abnormal hemoglobin, Hb S, which can distort red blood cells. In homozygous (bS/bS) individuals, altered hemoglobin (Hb) molecules precipitate inside the erythrocyte, changing its normal form into a sickle-shaped less functional cell. Homozygous individuals have a particularly increased risk of low-birth weight, thromboembolism and premature death due to early loss of splenic function or septic infection by encapsulated bacteria.<sup>19-20</sup> Although the occurrence of SCT varies greatly from state-to-state and among different races and ethnicities, every state and racial/ethnic population includes persons living with the condition. Definitive hemoglobin identification can be performed by protein sequencing, DNA analysis and HPLC combined with electrospray mass spectrometry in a specialized reference laboratory.<sup>21</sup> Such testing is indicated for infants with clinical or laboratory evidence of hemolysis or abnormal oxygen affinity and for infants without Hb A, especially if the unidentified variant is inherited with Hb S.<sup>22-23</sup> Identification of the hemoglobin variant to clarify genetic risks should also be considered in families where another hemoglobin abnormality (e.g. Hb S) is present.

### CONCLUSION

Studies suggested that there is need to maintain a primary prevention program to detect sickle cell disease at earlier stages despite having a large high risk. Preventive diagnosis and follow-up would reduce infant mortality by preventing the development of severe anemia as well as dangerous complications. In short, sickle cell disease surveillance would avert loss of life, measured as the number of years lost due to ill-health, disability or early death. Sequence analysis through Bioinformatics

approaches shows the mutations in HBB gene. Protein sequence analysis shows various domains in selected sequence. To study, sequence variations and to analyse these at molecular level can help to overcome many genetic disorders. We hope with some preliminary important actions and measures the early identification of sickle cell disease is possible in order to reduce health disparities in an already vulnerable population.

### ACKNOWLEDGEMENTS

The authors alone are responsible for the content and writing of this article. We are thankful to all authors for their support.

*Funding: No funding sources*

*Conflict of interest: None declared*

*Ethical approval: The study was approved by the Institutional Ethics Committee*

### REFERENCES

1. Silva WS, Lastra A, Oliveira SF, Guimarães NK, Grisolia CK. Avaliação da cobertura do programa de triagem neonatal de hemoglobinopatias em populações do Recôncavo Baiano, Brasil Cad Saúde Pública. 2006;22(12):2561-6.
2. Diniz D, Guedes C, Barbosa L, Tauil PL, Magalhães I. Prevalência do traço e da anemia falciforme em recém-nascidos do Distrito Federal, Brasil, 2004 a 2006 Cad Saúde Pública. 2009;25(1):188-94.
3. Sommer CK, Goldbeck AS, Wagner SC, Castro SM. Triagem neonatal para hemoglobinopatias: experiência de um ano na rede de saúde pública do Rio Grande do Sul, Brasil Cad Saúde Pública. 2006;22(8):1709-14.
4. Lobo CL, Bueno LM, Moura P, Ogeda LL, Castilho S, Carvalho SM. Triagem neonatal para hemoglobinopatias no Rio de Janeiro. Brasil Pan Am J Public Health. 2003;13(2/3):154-9.
5. Gelpi AP. Migrant populations and the diffusion of the sickle cell gene. Ann Intern Med. 1973;79:258-64.
6. Kulozik AE, Wainscoat JS, Serjeant GR, Kar BC, Al-Awamy B, Essan GJ et al. Geographical survey of beta S-globin gene haplotypes: evidence for an independent asian origin of sickle-cell mutation. Am J Hum Genet. 1986;39:239-44.
7. Stuart MJ, Nagel RL. Sickle-cell disease. Lancet. 2004;364:1343-60.
8. Chui DH, Dover GJ. Sickle cell disease: no longer a single gene disorder. Curr Opin Pediatr. 2001;13:22-7.
9. Lisowski L, Sadelain M. Current status of globin gene therapy for the treatment of  $\beta$ -thalassaemia. Br J Haematol. 2008;141(3):335-45.
10. Gelinas RE, Bender MA, Miller AD, Novak U. Long-term expression of the human  $\beta$ -globin gene after retroviral transfer into pluripotent

- hematopoietic stem cells of the mouse. *Adv Exp Med Biol.* 1989;271:135-48.
11. Gelinas RE, Bender MA, Miller AD, Novak U. Regulated expression of the human  $\beta$ -globin gene after retroviral transfer into murine and human hematopoietic cells. *Prog Clin Biol Res.* 1989;316B:235-49.
  12. May C, Rivella S, Callegari J, Heller G, Karen ML, Gaensler, Lucio Luzzatto. Therapeutic haemoglobin synthesis in  $\beta$ -thalassaemic mice expressing lentivirus-encoded human  $\beta$ -globin. *Nature.* 2000;406(6791):82-6.
  13. Pawliuk R, Westerman KA, Fabry ME, Payen E, Tighe R, Bouhassira EE. Correction of sickle cell disease in transgenic mouse models by gene therapy. *Science.* 2001;294(5550):2368-71.
  14. Pestina TI, Hargrove PW, Jay D, Gray JT, Boyd KM, Persons DA. Correction of murine sickle cell disease using  $\gamma$ -globin lentiviral vectors to mediate high-level expression of fetal hemoglobin. *Mol Ther.* 2008;17(2):245-52.
  15. Nagela RL, Bookchina RM, Johnsona J, Labieb D, Wajcmanb H, William A. Structural bases of the inhibitory effects of hemoglobin F and hemoglobin A2 on the polymerization of hemoglobin S. *Proc Natl Acad Sci USA.* 1979;76(2):670-2.
  16. Cavazzana-Calvo M, Payen E, Negre O, Wang G, Hehir K, Fusil F, Down J. Transfusion independence and HMGA2 activation after gene therapy of human  $\beta$ -thalassaemia. *Nature.* 2010;467(7313):318-22.
  17. World Health Organization. Sickle-cell anaemia. Report by the Secretariat. Fifty-ninth World Health Assembly. Provisional agenda item 11.4. A59/9. WHO. 2006. Last accessed on 2006.
  18. Powars DR, Chan LS, Hiti A, Ramicone E, Johnson C. Outcome of sickle cell anemia: A 4-decade observational study of 1056 patients. *Medicine.* Baltimore. 2005; 84:363-76.
  19. Brown AK, Sleeper LA, Miller ST, Pegelow CH, Gill FM, Waclawiw MA. Reference values and hematologic changes from birth to 5 years in patients with sickle cell disease. Cooperative Study of Sickle Cell Disease. *Arch Pediatr Adolesc Med.* 1994;148(8):796-804.
  20. Steinberg MH. Sickle cell anemia, the first molecular disease: overview of molecular etiology, pathophysiology and therapeutic approaches. *Scientific World Journal.* 2008;8:1295-324.
  21. Lane PA, Witkowska HE, Falick AM, Houston ML, McKinna JD. Hemoglobin D Ibadan - beta-0-thalassemia: detection by neonatal screening and confirmation by electrospray-ionization mass spectrometry. *Am J Hematol.* 1993;44:158-61.
  22. Strickland DK, Ware RE, Kinney TR. Pitfalls in newborn hemoglobinopathy screening: failure to detect beta-thalassemia. *J Pediatr.* 1995;127:304-8.
  23. Witkowska HE, Lubin BH, Beuzard Y, Baruchel S, Esseltine DW, Vichinsky EP. Sickle cell disease in a patient with sickle cell trait and compound heterozygosity for hemoglobin S and hemoglobin Quebec-Chori. *N Engl J Med.* 1991;325:1150-4.

**Cite this article as:** Mubeen H, Naqvi RZ, Masood A, Shoaib MW, Raza S. In silico mutation analysis of human beta globin gene in sickle cell disease patients. *Int J Res Med Sci* 2016;4:1673-7.