New Approach for Analysis and Prediction of Genetic Beta-Thalassemia Mutations Based On Bioinformatics Bioedit Tools.

Anar Auda Ablahad

Abstract: Beta-thalassemia is one of the most common single gene disorders affecting almost all the countries. It is a blood disorder that reduces the production of hemoglobin and commonley caused by inherited gene mutations of HBB gene copies from father and mother. Predicting the gene healthy plays essential role in saving human life. A new approach for the detection of beta-thalassemia mutations via checking and testing the gene healthy and find if the gene hold mutation will be apply based on bioinformatics tools via bioedit package. This method allows genotyping of the HBB gene of patient's DNA. The system provides a bases for speed (2 seconds only), rapid, simple, and reliable detection of the numerous known beta-thalassemia gene mutations.

Index Terms— Beta-thalassemias; HBB gene equence; Sequence Alignment; FASTA; ClustalW; Bioinformatics Tools; gene mutations.

1 INTRODUCTION

Thalassemia is one of the major hemoglobin- pathies among the population all around the world [1]. it considered as the most widespread genetic mutation. According to the World Health Organization (WHO) between 1.5-7% of the world population are carriers for this disease, and every year 60,000-400,000 birth of new patients are reported. In Israel, the incidence of carriers for B-thalassemia is around 20% among the Jewish from Kurdish origin and around 5-10% among the Arab population. And every year 60,000-400,000 birth of new patients are reported [2]. It is inherited blood disorders which passed from parents to children through genes and has life-long implications for patients and families. Thalassemia cause the body to make fewer healthy red blood cells and less hemoglobin than normal [3].

Beta thalassemia is a blood disorder that reduces the production of hemoglobin. Hemoglobin is the iron-containing protein in red blood cells that carries oxygen to cells throughout the body. In people with beta thalassemia, low levels of hemoglobin lead to a lack of oxygen in many parts of the body. Affected individuals also have a shortage of red blood cells (anemia), which can cause pale skin, weakness, fatigue, and more serious complications. People with beta thalassemia are at an increased risk of developing abnormal blood clots [4].

Because thalassemia is inherited from parents to child in genes, sometimes changes occur to genes, resulting in medical conditions. Such as changes occur to beta globin genes in beta thalassemia: A person normally inherits two globin genes for the production of the beta globin protein in hemoglobin. A person may have an alteration (mutation) in one of their two globin genes. This person is called a carrier of Thalassemia and is healthy. Carriers may be at risk of having a child affected with beta B thalassemia major if their partner is also a carrier of B thalassemia. When a person has alterations (mutations) in both of their B globin genes, they have a severe condition called B thalassemia major. Thalassemia major results in severe anemia requiring lifelong treatment [5].

B-thalassemia major can becaused by homozygosis or compound heterozygosis for B-globin gene mutations (HBB gene)[4]. The HBB gene provides instructions for making a protein called beta-globin. Beta-globin is a component (subunit) of a larger protein called hemoglobin, which is located inside red blood cells. Without proper amounts of beta-globin, sufficient hemoglobin cannot be formed. A lack of hemoglobin disrupts the normal development of red blood cells. A shortage of mature red blood cells prevents these cells from carrying and delivering enough oxygen to satisfy the body's energy needs. A lack of oxygen in the body's tissues can lead to poor growth, organ damage, and other health problems associated with beta thalassemia [4].

The HBB gene encodes an important blood protein called beta globin. A person with beta-thalassemia carries a mutation in both copies of the HBB gene, completely halting production of the beta globin protein. Without beta globin, the important oxygen-carrying protein, hemoglobin, can not be made. Although oxygen can be carried by a less efficient form of hemoglobin, most of the affected red blood cells die [6].

Most cases are inherited from parents who both have diseased alleles of the HBB gene which is located on from base pair 5,246,695 to base pair 5,248,300 on chromosome 11.

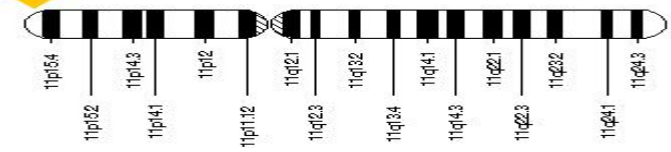


Figure 1: HBB Location on chromosome 11.

The beta-thalassemias were among the first human genetic diseases to be examined by means of new techniques of recombinant DNA analysis. In general, the molecular pathology of disorders resulting from mutations in the nonalpha-globin gene region is the best known [7].

2. RELATED WORK

Fettah A., et.al [8], Analyzed and tested for 106 turkey patients gene sequences for B-globin gene mutations using DNA analysis. And classified as holding B-thalassemia major or B-thalassemia intermedia based on their age at diagnosis. The result showed various types of mutations types in gene regions and passed them to turkey hospital for future gene tests.

Atanasovska B, et.al [9], Proposed approach applicable in a range of Mediterranean countries, they offered a combination of high accuracy and rapidity exploiting standard techniques and widely available equipment. Beta globin Detection further adapted to particular populations by including/excluding assayed mutations. And facilitate future modifications by providing detailed information on assay design

Verma IC, et.al [10]l, characterized the mutations in 1050 carriers of the beta-thalassemia gene and analyzed their regional distribution in India. The majority of beta-thalassemia carriers were migrants from Pakistan and their pattern of mutations differed from the rest. The paper result helped to successfully establish a program of genetic counselling and prenatal diagnosis of beta-thalassemia in order to reduce the burden of this disease in India.

P. Lahiry, et.al [11], suggested that "efforts to more completely characterize the HBB mutation distribution in high-risk areas, such as the Indian Subcontinent and the Middle East" may lead to improved diagnosis with earlier and more effective intervention strategies. The concluded that beta-thalassemia is highly prevalent and is a major public health problem in the malaria endemic areas of the Indian

The weakness of all these methods and techniques they focused in diagnostic or detecting Thalassemia were based on genes mutations locally. Means that comparison of patients' gene was made to conclude specific annotations and mutations in the selected country. Without taking in consideration the abnormality of gene state wheather normal or abnormal.

The motivation behind this new approach of beta thalassemia mutations detection is to check the HBB state as it is the gene that mutations on it will causes Beta thalassemia. With taking some features and helpful information about the gene.

3 THE PROPOSED APPROACH.

The main task and aims of this proposed new approach is to diagnostic and detecting mutations in the Beta-globin gene(HBB gene) sequence by comparisions between selected parents gene equences wit the normal gene (the gene without annotation). In order to check the normality of gene state. The new approach contains many importatnt steps to reach the goal of detection as follows

1. Start
2. Select the gene causes to the genetic disease
3. Find the reference gene sequence (gene without mutations)
4. Get the patient's gene sequence for detection.
5. Make fasta file of the two sequences.
6. Use ClustalW alignment tool to find similarity.
7. if changes occure, then
8. Follow and monitor transicription and translation processes.
9. If the protein sequences result with changes between the sequences. The,
10. The gene sequence is infected and the patient has Beta thalassemia
11. End

A. BIOINFORMATIC DATABASES.

because the environment plays important role in the DNA sequences shapes, it is important to find the reference gene sequence that doent hold any annotations or mutation so as to adobt it for alignment and classification. Bioinformatic databases provide useful references to find and deal with gene sequences by providing GenBank sequences with Reference genome that help any one to make sure of the gene healthy.

B. BIOINFORMATICS TOOLS AND PROCESSES.

1. FASTA: For every bioinformatics tools selected to complete tasks. FASTA format is required to be set as a standard expression of the entire file to the elected bioinformatics tool. FASTA file could also help in find a sequence of the required gene agenist keywords.

2. ClustalW: it is a powerful technique for checkup the sequences similarity whether the patient has malignant mutation related to ubnormal disorder or no. by using it within BioEdit package; it can also provide the way to monitor the processes of transcription and translation to get protein sequences that controll the function in gene. In this new approach, ClustalW accept FATA file that contains the normal gene (HBB gene sequence) and the patient/ parents HBB gene in order to check the gene's healthy as follows.

• *Anar Auda Ablahad is assistant lecturer at University Of Duhok-Electrical and computer eng department-iraqt. She is graduated from Software Engineering collage at mosul university -Iraq. And got master degree in computer scinces at Zakho University. Her field interest in programming, networking, Bioinformatic, Neural network. E-mail: anar_alkasyonan@yahoo.com*

3. BLAST, the basic local alignment search tool is the comon widely used bioinformatics tools to searching for similarities between biological sequences which performs comparisons between pairs of sequences, searching for regions of local similarity to start sequence analysis.

1. Start
2. Make Fasta File (Normal& Person's) HBB Gene
3. Use The File as Input to BioEdit ClustalW
4. Check Neocluotide alignment Of The Entire Sequences.
5. Seq1 !=Seq2?
6. Convert Sequences To protiens.
7. Use Alignment Tool(Pairwise Alignment) for Similarity Check.
8. Differences Found? (Seq1 !=Seq2)? Then
9. HBB Gene is at Risk and it is Candidate for Thalassemia Disease.
10. End

By reaching protien checks, we can make sure of the basic role which is (**A person with beta-thalassemia carries a mutation in both copies of the HBB gene, completely halting production of the beta globin protein**).

4. Expermintal Results

The quality of medical care has always been a key issue for both practitioners and patients and the highest standards and practice guidelines are expected in all fields of medicine. The diagnosis of thalassemia is often important because of atypical clinical histories

The Results fron implementing the new approach based on HBB gene which is the only gene causes B-Thalassemia can be implemented in many step as:

1. Obtainig the normal HBB gene sequence from official bioinformatics databases according to the suitable invironment . in this system, the normal gene adobted from NCBI reference genome databases as follows:

NCBI – > Gene-> Reference genome-> HBB gene sequence

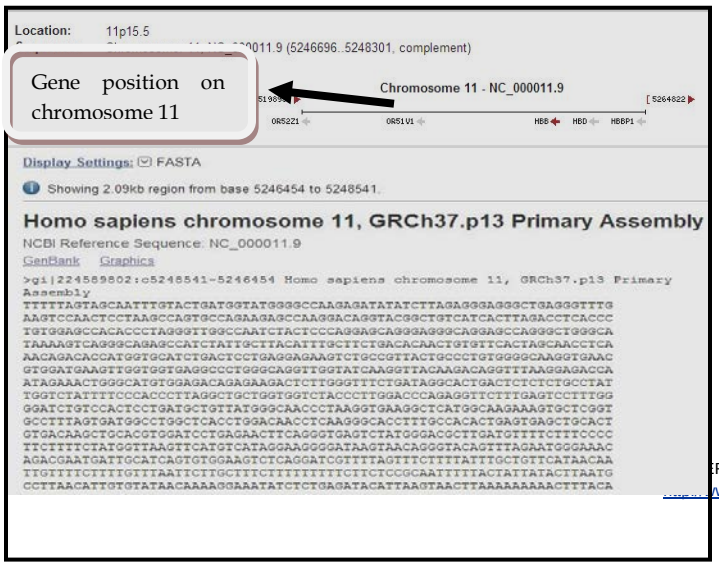


Figure 2: Obtaining the normal gene from NCBI.

2. in order to analyze and test the HBB gene sequence of the patient to extract the similar sequences reported in the world to make sure of the gene healthy. We can use blast suite tool of bioinformatics tool to handle the analysis process. The HBB gene.

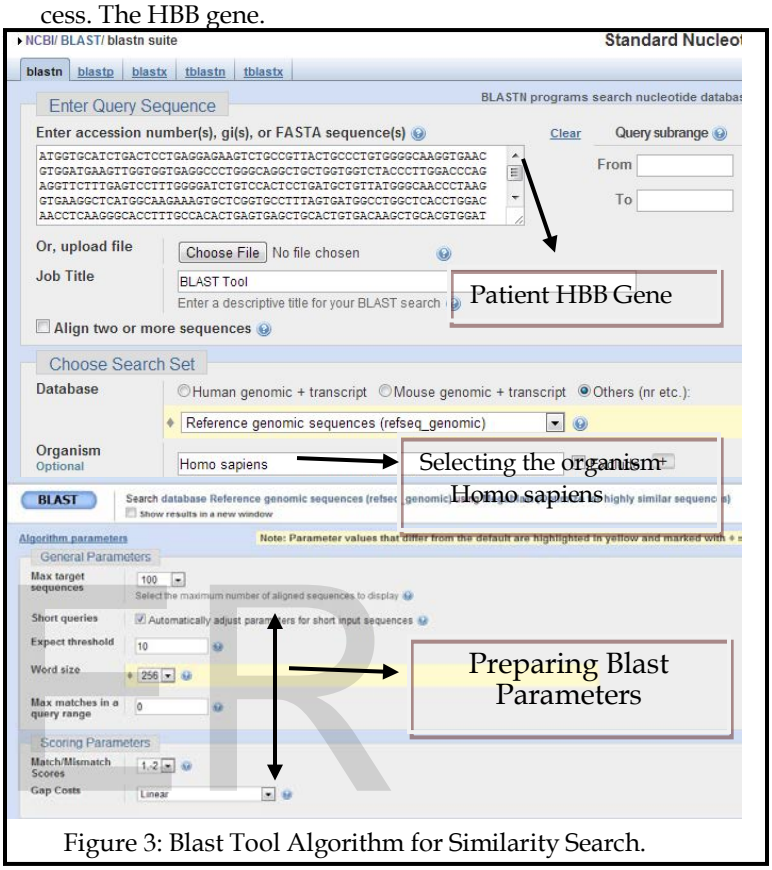


Figure 3: Blast Tool Algorithm for Similarity Search.

The blast results will appear after a while with all the related sequences at NCBI databases.

3. The important step beigns in this stage after analysis and search. In this stage the normal HBB gene sequence and patient gene sequence is used as FASTA file input to the bioedit package tool to test the patient's gene if its healthy or hold somatic mutations that causes thalassemia.

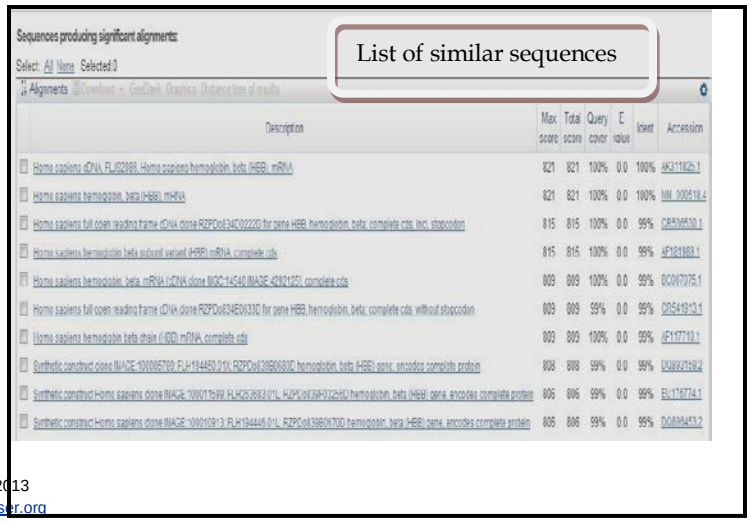
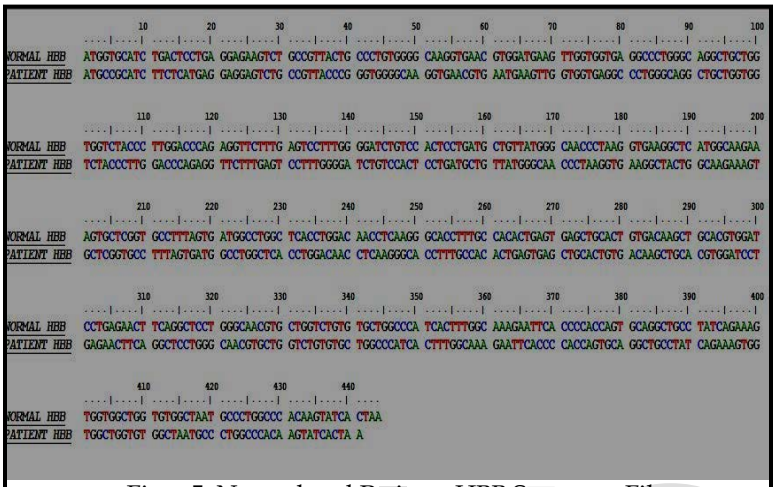


Figure 4: Blast Results Show all Similar Sequences at NCBI.

3. The important step beigns in this stage after analysis and search. In this stage the normal HBB gene sequence and patient gene sequence is used as FASTA file input to the bioedit package tool to test the patient's gene if its healthy or hold somatic mutations that causes thalassemia.



Figre 5: Normal and Patients HBB Sequence File.

First the clustalW tool is used for alignment of the two HBB sequences at Neocleotide level.

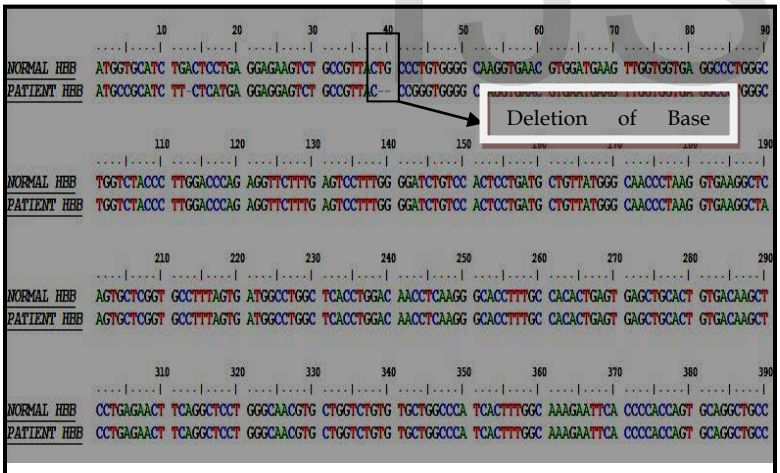


Figure 6: Shows the Alignment Result (Deletion Mutation).

In this stage and because the result shows mutation in some positions on patient's gene. This means that there is a chance of having Beta thalassemia. In order to make sure of this result the two sequences should convert to protein via DNA-RNA-Protien sequences to find if the genechanges its function. If the protein sequence keeps the sequences similar 100% this mean that the gene is healthy, otherwise the gene would be at highly risk of Beta thalassemia disease.

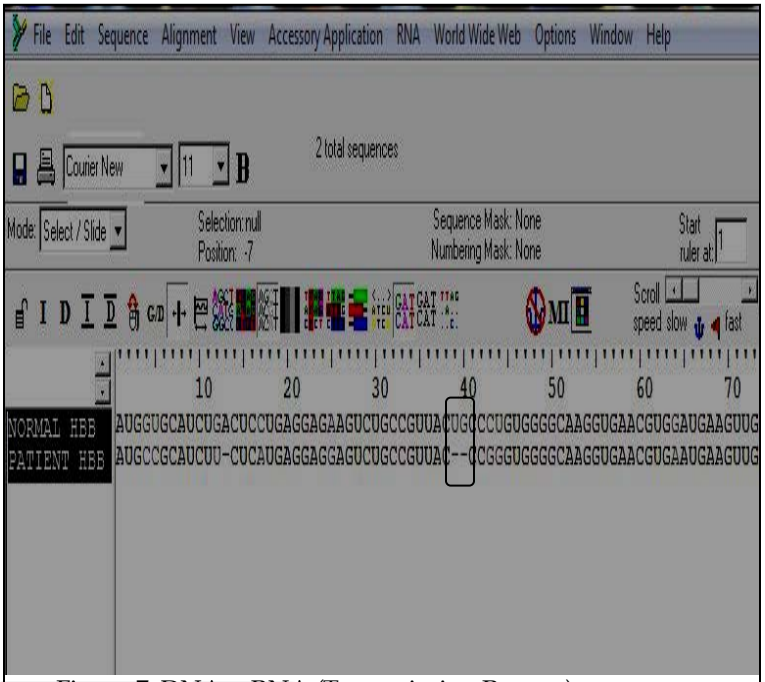


Figure 7: DNA-> RNA (Transcription Process).

The final stage is the protein sequence convertion (transla tion process). if differences occure a gain then the gene is ub normal and the disease is actually happened.

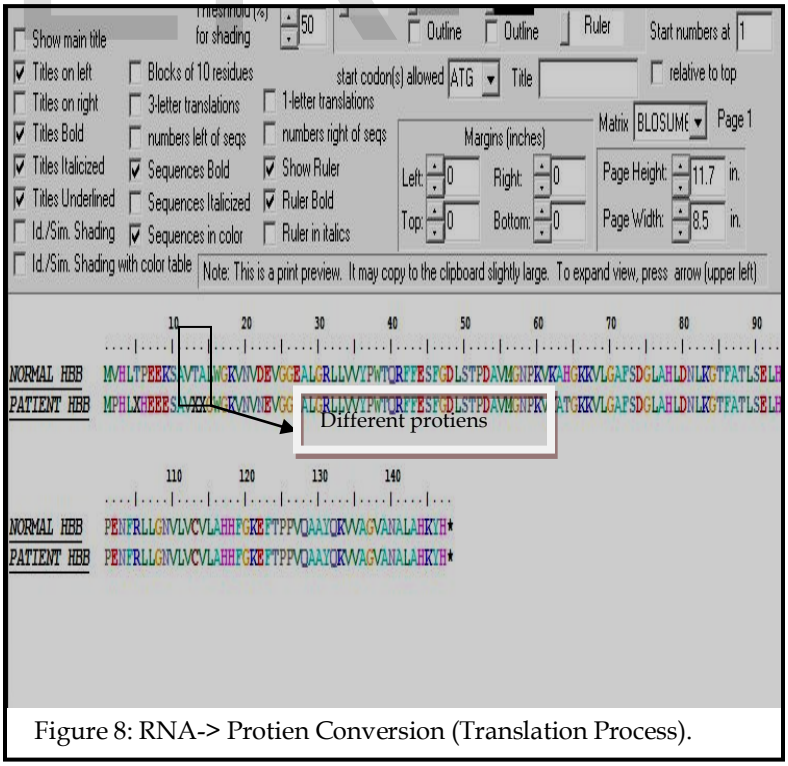


Figure 8: RNA-> Protien Conversion (Translation Process).

The result shows that the unlike protein sequences alignment de-scribe the ubnormality of the HBB gene and the gene finally hold

Beta thalassemia disease.

Discussion of the Results

Comparing the results of current new approach with previous works can be list as:

Features	The New Approach	Fattah A. et.al	Verma IC, et.al
The Aims	Analysis, alignment and mutations check and gene healthy test.	tested for 106 turkey patients gene sequences for B-globin gene mutations	characterized mutations in 1050 carriers of B-thalassemia gene and analyzed their distribution in India
The approach	Used DNA (both Nucleotide & protein)	Used DNA (Nucleotide only)	Used DNA (Nucleotide only)
General method	Yes	No-limited	No-limited
The results can support	Bioinformatics, molecular biology, doctors	Molecular& doctors	Doctors

Table 1: Comparision with Previous Works.

CONCLUSION

The main conclusions which obtained from implementing the proposed novel method of disease prediction are:

1. This proposed approachl is first prediction method gives accurate results, because it is not based on malignant mutations of genes which caused the disease only, but also base on their proteins.
2. This new approach suggested a general prediction method based on mutational in genes caused the disease, i.e. can implement this novel method for any disease when the mutations of its gene which caused the disease are known.
3. Offers an automatic, cost effective and friendly diagnosis system for detecting malignant mutation Beta thalassemia as shown in Table 1. it can use by any researcher or patient who needed to test malignant mutations at genes which caused breast any type of thalassemia.

REFERENCES

[1] Amrita Panja , Tapan K Ghosh , Anupam Basu, "Genetics of Thalassemia in Indian Population", Journal of Community Nutrition & Health, Vol.1, Issue 1, 2012.

[2] Dr Koren A., HaEmek Medical Center , "Detection of β Thalassemia Carriers by Red Cell Parameters Obtained From the H2 Automatic Counter, clinical trails.gov. , July 2013

[3] Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, National Institutes of Health, "Genetic Home Reference", 2013.

[4] Mercy Hospital for Women, Monash Medical Centre, "Beta (B) thalassaemia", © ThalAust Inc. and Southern Health, 2005-2009.

[5] D.S. Coming and O.G. Staadt, "Velocity-Aligned Discrete Oriented Polytopes for Dynamic Collision Detection," *IEEE Trans. Visualization and Computer Graphics*, vol. 14, no. 1, pp. 1-12, Jan/Feb 2008, doi:10.1109/TVCG.2007.70405. (IEEE Transactions)

[6] Naga Ganesh Balne and C.S.V. Ramachandra Rao, "Role of XmnI restriction site polymorphism and JAK2 gene mutation in β-Thalassemia", International Research Journal of Biological Sciences ISSN 2278-3202, Vol. 2(1), 41-45, January (2013)

Fettah A, Bayram C, Yarali N, Isik P, Kara A, Culha V, Tunc B., "Beta-globin Gene Mutations in Turkish Children with Beta-Thalassemia: Results from a Single Center Study., Mediterr J Hematol Infect Dis. 2013 Sep 2;5(1):e2013055. doi: 10.4084/MJHID.2013.055.

[9] Atanasovska B, Bozhinovski G, Plaseska-Karanfilska D, Chakalova L (2012) Efficient Detection of Mediterranean β-Thalassemia Mutations by Multiplex Single-Nucleotide Primer Extension. PLoS ONE 7(10): e48167. doi:10.1371/journal.pone.0048167.

[10] Verma IC, Saxena R, Thomas E, Jain PK. "Regional distribution of beta-thalassemia mutations in India", PMID: 9225979. [PubMed - indexed for MEDLINE].

[11] P Lahiry, S A Al-Attar, R A Hegele, "Understanding Beta-Thalassemia with Focus on the Indian Subcontinent and the Middle East", The Open Hematology Journal, 2008, 2, 5-13.

IJSER