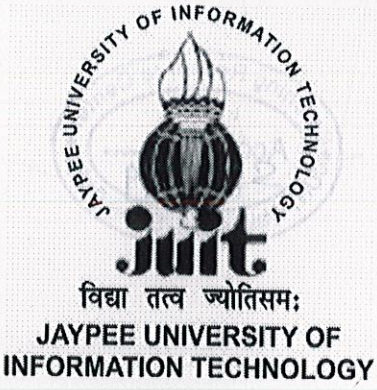
Single Nucleotide Polymorphism Survey for Beta Thalassemia in the Population from Dehradun and Varanasi

By

Amit Kumar Singh (061703)
Prateek Tripathi (061577)

under the Supervision of
Dr. Harvinder Singh



MAY2010

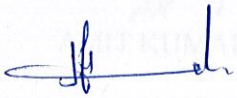
Submitted in partial fulfillment of the Degree of Bachelor of Technology
In Biotechnology

DEPARTMENT OF
BIOTECHNOLOGY & BIOINFORMATICS
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT, SOLAN, HP, INDIA

CERTIFICATE

This is to certify that the work entitled, "**SINGLE NUCLEOTIDE POLYMORPHISM SURVEY FOR BETA THALASSEMIA IN THE POPULATION FROM DEHRADUN AND VARANASI**" submitted by AMIT KUMAR SINGH and PRATEEK TRIPATHI. In partial fulfillment for the award of degree of Bachelor of Technology of Jaypee University of Information Technology has been carried under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

SUPERVISOR:



Dr. Harvinder Singh

ACKNOWLEDGEMENT

We hereby acknowledge with gratitude for the co-operation and help given to us by allthe member of this organization (Jaypee University of Information Technology) in completing the final year project. With proud privilege and profound sense of graditude, we acknowledge our indebtedness to our project guide **Dr. Harvinder singh** for his valuable guidance inputs, suggestions, constant encouragement and co-operation. We express our heartfelt thanks to our Head of Department **Dr. R.S.Chauhan** for providing us with the opportunities of duing this project. We express our gratitude to all our lab Attendents of Dept. of Biotechnology and Bioinformatic Jaypee University of Information Technology, Solan for their numerous help. Last, but not the least, we express our thankfulness to our parents, brothers andSister whose support, love, affection has been a source of encouragement which always motivate us to move ahead in our life.

AMIT KUMAR SINGH

061703

PRATEEK TRIPATHI

061577

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PCR -Polymerase Chain Reaction

SNP - Single nucleotide polymorphism

RFLP -Restriction Fragment Length polymorphism

CAPS -Cleaved Amplified Polymorphic sequences

DNA -Deoxy Ribonucleic Acid

RE -Restriction Enzyme

RBCs -Red Blood Cells

ABSTRACT

β-THALASSEMIA is a group of an inherited disease characterized by an abnormality of β globin production. More than 90 types of mutation have been reported. In general, each population has a different group of mutations, consisting of a few very common ones and a variable number of rare ones and also we are trying to find out whether there is any mutation in heme part also or not or it is complete mutation in globin chain only.

The study of these genetic mutations has been based on polymerase chain reaction (PCR) and visualization of the fragments after electrophoresis in agarose gel.

Traditional detection method such as low mean corpuscular volume, etc has relied on the hematologic features, which are unable to differentiate between iron deficiency anemia and mild β-thalassemia. The reliability of diagnosis can be increased by employing DNA-based techniques such as PCR. The project aims at discovering new SNP by sequencing of the mutant samples. Then developing multiplexing primers for easy diagnosis.

CHAPTER 1
INTRODUCTION

1.1 THALASSEMIA :

Thalassemia is an inherited autosomal recessive blood disease. In thalassemia, the genetic defect results in reduced rate of synthesis of one of the globin chains that make up hemoglobin. Reduced synthesis of one of the globin chains can cause the formation of abnormal hemoglobin molecules, thus causing anemia, the characteristic presenting symptom of thalassemias.

The thalassemia trait may confer a degree of protection against malaria, which is or was prevalent in regions where the trait is common, thus conferring a selective survival advantage on carriers and perpetuating the mutation. In that respect the various thalassemias resemble another genetic disorder affecting hemoglobin, sickle cell disease.

Medical therapyfor beta thalassemia primarily involves iron chelation. Deferoxamine is the intravenously administered chelation agent currently approved for use in United States. Diferasirox is an oral iron chelation drug also approved in the US in 2005.

The antioxidant indicaxanthin, found in beets, in a spectrophotometric study showed that indicaxanthin can reduce perferryl-Hb generated in solution from met-Hb and hydrogen peroxide, more effectively than either Trolox or Vitamin C.

Thalassemia causes the body to make fewer healthy red blood cells and less hemoglobin than normal. Hemoglobin is an iron-rich protein in red blood cells. It carries oxygen to all parts of the body. It also carries carbon dioxide (a waste gas) from the body to the lungs, where it's exhaled.

People who have thalassemia can have mild or severe anemia. This condition is caused by a lower than normal number of red blood cells or not enough hemoglobin in the red blood cells.

1.2 THALASSEMIA AND DIAGNOSIS

Thalassemia is the most common, inherited single gene disorder in the world. Early comprehensive treatment has changed thalassemia from a fatal pediatric disease to one in which patients live productive lives throughout adulthood. Advances in treatment are exciting, resulting in the potential for cure and improved quality of life. However, many patients never receive the information needed to make educated decisions about treatment. Many ethnic groups are unaware of their genetic risk of thalassemia and do not receive genetic counseling or pre-natal diagnosis. Others have heard of pre-implantation genetic diagnosis as a technique to achieve a healthy newborn, but do not have access to objective medical information of its risks and benefits.

1.3 THALASSEMIA AND HAPLOTYPING

A limited number of haplotypes are found in each population, and each haplotype is usually associated with one specific type of thalassemia.. Haplotype analysis is useful in identifying a thal chromosome, with an accuracy of about 90%. The aim of the present project is to do the haplotype analysis for -thalassemia chromosomes among minor and major thalassemia patients from the North INDIA and to compare them with the haplotype background of normal individuals in the region.

1.4 POLYMERASE CHAIN REACTION

Polymerase chain reaction (PCR) is a technique to amplify a single or few copies of a piece of DNA across several orders of magnitude, generating millions or more copies of a particular DNA sequence. The method relies on thermal cycling, consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA. Primers containing sequences complementary to the target region along with a DNA polymerase are key components to enable selective and repeated amplification. As PCR progresses, the DNA generated is itself used as a template for replication, setting

in motion a chain reaction in which the DNA template is exponentially amplified. PCR can be extensively modified to perform a wide array of genetic manipulations.

1.5 SINGLE NUCLEOTIDE POLYMORPHISM

A **Single Nucleotide Polymorphism** (SNP) is a DNA sequence variation occurring when a Single nucleotide – A, T, G, or C – in the genome (or other shared sequence) differs between members of species (or between paired chromosomes in an individual).

The following objectives were set for the present study :

- To study the distribution of major β-thalassemia mutations in different parts of India with emphasis on Dehradun and Varanasi.
- Designing primers from the flanking regions of target SNPs present in beta globin gene.
- Sequencing and identification of de novo SNPs.
- To study the set of mutation that is in SNP and calculate O_2 binding affinity of β-globin chain of hemoglobin.

CHAPTER 2
REVIEW LITERATURE

2.1 HAPLOTYPE :

In genetics, a haplotype (from the Greek : "one fold, single, simple") is a combination of alleles at multiple loci that are transmitted together on the same chromosome. Haplotype may refer to as few as two loci or to an entire chromosome depending on the number of recombination events that have occurred between a given set of loci.

In a second meaning, haplotype is a set of single nucleotide polymorphisms (SNPs) on a single chromatid that are statistically associated. It is thought that these associations, and the identification of a few alleles of a haplotype block, can unambiguously identify all other polymorphic sites in its region. Such information is very valuable for investigating

From SNP to Haplotype

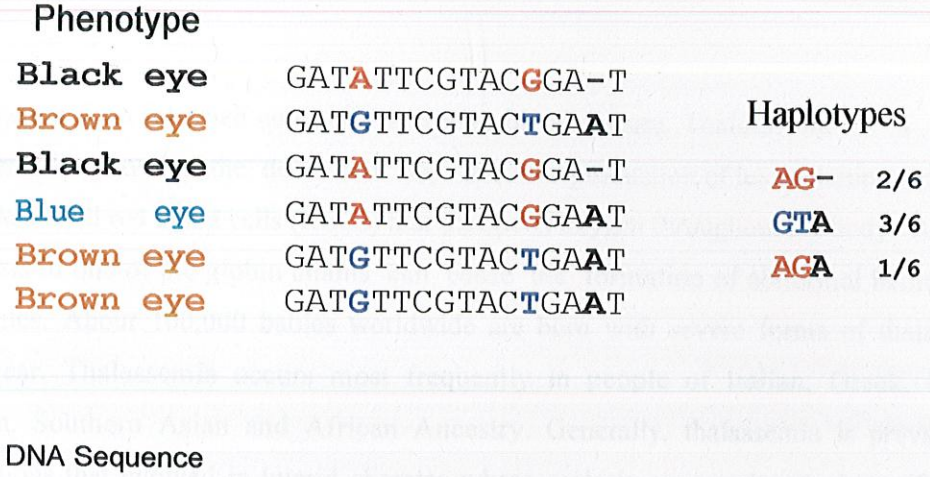


FIG .2.1 From SNP to Haplotype

2.2 THALASSEMIA:

Thalassemia is an inherited autosomal recessive blood disease. Thalassemia is a genetic disorders that involve the decreased and defective production of hemoglobin, a molecule found inside all red blood cells (RBCs) that transports oxygen throughout the body. Reduced synthesis of one of the globin chains can cause the formation of abnormal hemoglobin molecules. About 100,000 babies worldwide are born with severe forms of thalassemia each year. Thalassemia occurs most frequently in people of Italian, Greek, Middle Eastern, Southern Asian and African Ancestry. Generally, thalassemia is prevalent in populations that evolved in humid climates where malaria was endemic, but affects all races. Thalassemia is particularly associated with Arabs, people of Mediterranean origin, and Asians. Thalassemia occurs because the body doesn't make enough healthy red blood cells and hemoglobin. The severity of symptoms depends on the severity of the disorder.

➢ *2.2.1 No Symptoms*

Alpha thalassemia silent carriers generally have no signs or symptoms of the disorder. This is because the lack of alpha globin protein is so small that hemoglobin works normally.

➢ *2.2.2 Mild Anemia*

People who have alpha or beta thalassemia trait can have mild anemia. However, many people with this type of thalassemia have no signs or symptoms. Mild anemia can make you feel tired. It's often mistaken for iron-deficiency anemia.

➢ *2.2.3 Mild to Moderate Anemia and Other Signs and Symptoms*

People with beta thalassemia intermedia have mild to moderate anemia. They also may have other health problems, such as:

Slowed growth and delayed puberty. Anemia can slow down a child's growth and development.

Bone problems : Thalassemia may make bone marrow (the spongy material inside bones that makes blood cells) expand. This causes wider bones than normal. Bones also may be brittle and break easily.

An enlarged spleen: The spleen is an organ that helps your body fight infection and removes unwanted material. When a person has a thalassemia, the spleen has to work very hard. As a result, the spleen becomes larger than normal. This makes anemia worse. If the spleen becomes too large, it must be removed.

2.3 PATHOPHYSIOLOGY

Hemoglobin, the oxygen-carrying protein in red blood cells, is made up of two chains, an alpha chain and a beta chain. These two chains are made from specific genes we inherit from our parents. When these specific genes are not working properly, hemoglobin production is affected. There are two major types of thalassemia.

In α thalassemia, production of α globin chain is affected, while in β thalassemia production of the β globin chain is affected.

2.3.1 α thalassemia :

The α thalassemia involve the genes HBA1 and HBA2, inherited in a Mendelian recessive fashion. It is also connected to the deletion of the 16p chromosome. α thalassemia result in decreased alpha-globin production, therefore fewer alpha-globin chains are produced, resulting in an excess of β chains in adults and excess γ chains in newborns. The excess β chains form unstable tetramers

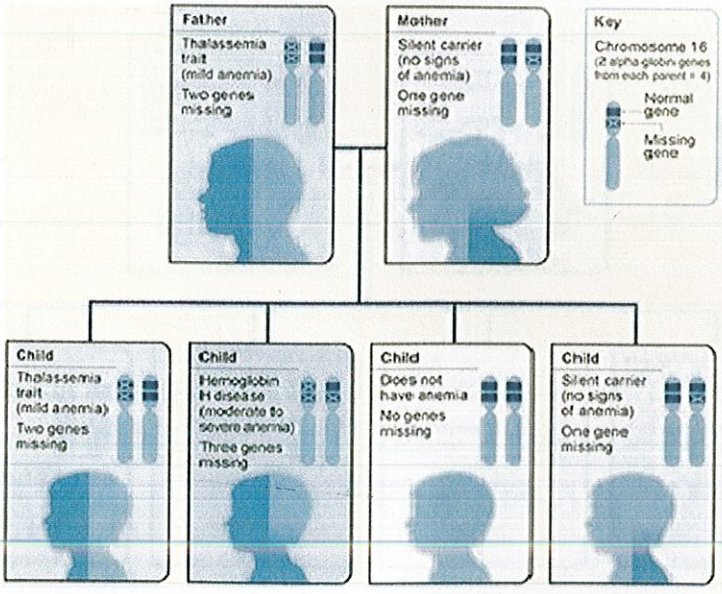


FIG 2.2. Inheritance Pattern for Alpha Thalassemia

The diagram shows one example of how alpha thalassemia is inherited. The alpha globin genes are located on chromosome 16. A child inherits four alpha globin genes — two from each parent. In this example, the father is missing two alpha globin genes and the mother is missing one alpha globin gene.

Therefore, each child has a 25 percent chance of inheriting two missing genes and two normal genes (thalassemia trait), three missing genes and one normal gene (hemoglobin H disease), four normal genes (no anemia), or one missing gene and three normal gene.

2.3.2 Beta (β) thalassemia

Beta thalassemia is due to mutations in the HBB gene on chromosome 11. The severity of the disease depends on the nature of the mutation.

In either case there is a relative excess of α chains, but these do not form tetramers *Any given individual has two β globin alleles.*

If only *one* β globin allele bears a mutation, the disease is called β thalassemia minor. If *both* alleles have thalassemia mutations, the disease is called β thalassemia Major

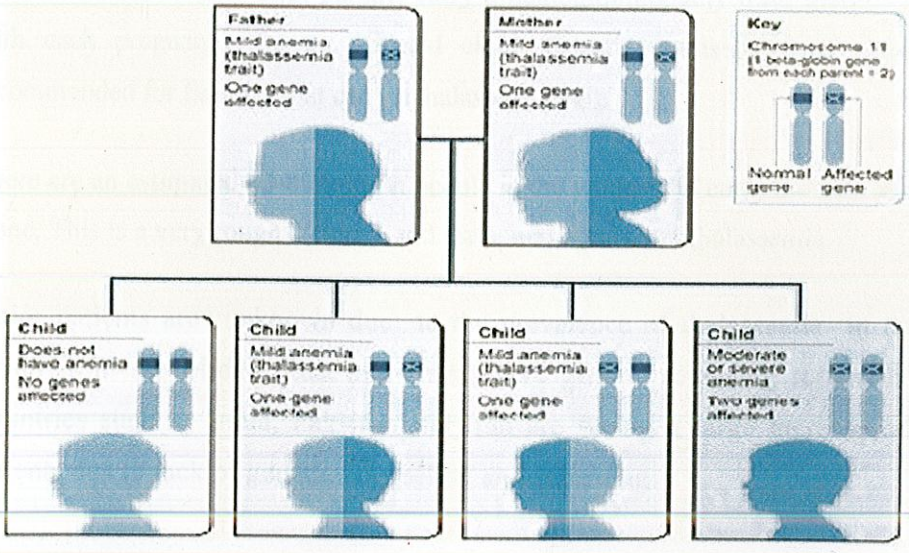


FIG 2.3. Inheritance Pattern for Beta Thalassemia

The diagram shows one example of how beta thalassemia is inherited. The beta globin gene is located on chromosome 11. A child inherits two beta globin genes —one from each parent. In this example, each parent has one altered beta globin gene. Therefore, each child has a 25 percent chance of inheriting two normal genes (no anemia), a 50 percent chance of inheriting one altered gene and one normal gene (beta thalassemia trait), or a 25 percent chance of inheriting two altered genes (beta thalassemia major).

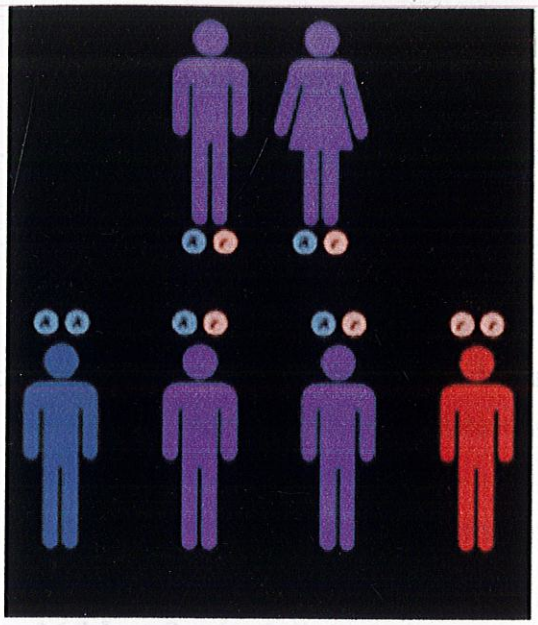
2.4 GENETIC PREVALENCE

α and β thalassemia are often inherited in an autosomal recessive fashion although this is not always the case. Cases of dominantly inherited α and β thalassemia has been reported, the first of which was in an Irish family who had a two deletions of 4 and 11 bp in exon 3 interrupted by an insertion of 5 bp in the β-globin gene.

For the autosomal recessive forms of the disease both parents must be carriers in order for a child to be affected. If both parents carry a hemoglobinopathy trait, there is a 25% chance with each pregnancy for an affected child. Genetic counseling and genetic testing is recommended for families that carry a thalassemia trait.

There are an estimated 60-80 million people in the world who carry the beta thalassemia trait alone. This is a very rough estimate and the actual number of thalassemia

Major patients are unknown due to the prevalence of thalassemia in less developed countries in the Middle East and Asia where genetic screening resources are limited. Countries such as India, Pakistan and Iran are seeing a large increase of thalassemia patients due to lack of genetic counseling and screening.



Thalassemia has an autosomal recessive pattern of inheritance

FIG 2.4. Pattern Of Inheritance

2.5 CAUSES OF THALASSEMIA

The cause of thalassemia is defects in the genes that make hemoglobin. The only way to get thalassemia is to inherit one or more defective hemoglobin genes from your parents.

Hemoglobin is a red, iron-rich protein found in red blood cells. Hemoglobin enables red blood cells to carry oxygen from your lungs to all parts of your body and to carry carbon dioxide from other parts of your body to your lungs so that it can be exhaled. Most blood cells, including red blood cells, are produced regularly in your bone marrow — a red, spongy material found within the cavities of many of your large bones.

Thalassemia disrupts the normal production of hemoglobin and leads to a low level of

hemoglobin and a high rate of red blood cell destruction, causing anemia. When you're anemic, your blood doesn't have enough red blood cells to carry oxygen to your tissues — leaving you fatigued.

2.6 TREATMENT OF THALASSEMIA

2.6.1 Blood Transfusions. The most common treatment for all major forms of thalassemia is red blood cell transfusions. These transfusions are necessary to provide the patient with a temporary supply of healthy red blood cells with normal hemoglobin capable of carrying the oxygen that the patient's body needs. While thalassemia patients were given infrequent transfusions in the past, clinical research led to a more frequent program of regular blood cell transfusions that have greatly improved the patients' quality of life. Today, most patients with a major form of thalassemia receive red blood cell transfusions every two to three weeks, amounting to as much as 52 pints of blood a year.

2.6.2 Iron Overload Because there is no natural way for the body to eliminate iron, the iron in the transfused blood cells builds up in a condition known as "iron overload" and becomes toxic to tissues and organs, particularly the liver and heart. Iron overload typically results in the patient's early death from organ failure.

2.6.3 Chelation Therapy To help remove excess iron, patients undergo the difficult and painful infusion of a drug, Desferal. A needle is attached to a small battery-operated infusion pump and worn under the skin of the stomach or legs five to seven times a week for up to twelve hours. Desferal binds iron in a process called "chelation." Chelated iron is later eliminated, reducing the amount of stored iron.

2.7 DIAGNOSIS

Doctors diagnose thalassemia using blood tests, including a complete blood count (CBC) and special hemoglobin tests. Tests on the amount of iron in the blood to find out whether the anemia is due to iron deficiency or thalassemia.

If a person has thalassemia, blood tests may reveal:

- A low level of red blood cells
- Smaller than expected red blood cells
- Pale red blood cells
- Red blood cells that are varied in size and shape
- Red blood cells with an uneven hemoglobin distribution, which gives the cells a bull's-eye appearance under the microscope

Blood tests may also be used to:

- Measure the amount of iron in a person's blood
- Evaluate his or her hemoglobin
- Perform DNA analysis to diagnose thalassemia or to determine if a person is carrying defective hemoglobin genes

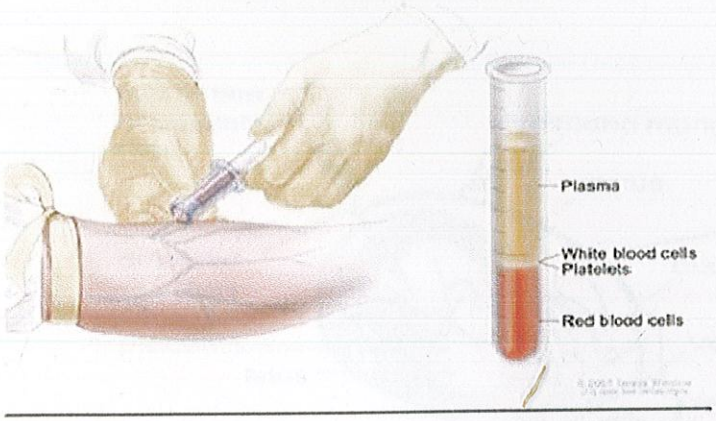


FIG 2.5. Complete blood count (CBC). Blood is collected by inserting a needle into a vein and allowing the blood to flow into a tube. The blood sample is sent to the laboratory and the red blood cells, white blood cells, and platelets are counted.

Prenatal testing

Testing can be done before a baby is born to find out if he or she has thalassemia and determine how severe it may be. Tests used to diagnose thalassemia in unborn babies include :

- *Chorionic villus sampling*. This test is usually done around the 11 th week of pregnancy and involves removing a tiny piece of the placenta for evaluation.
- *Amniocentesis* This test is usually done around the 16th week of pregnancy and involves taking a sample of the fluid that surrounds the baby.

- *Fetal blood sampling*. This test can be performed after 18 weeks of gestation and involves taking blood from the fetus or the blood vessels in the umbilical cord using an ultrasound-guided needle.

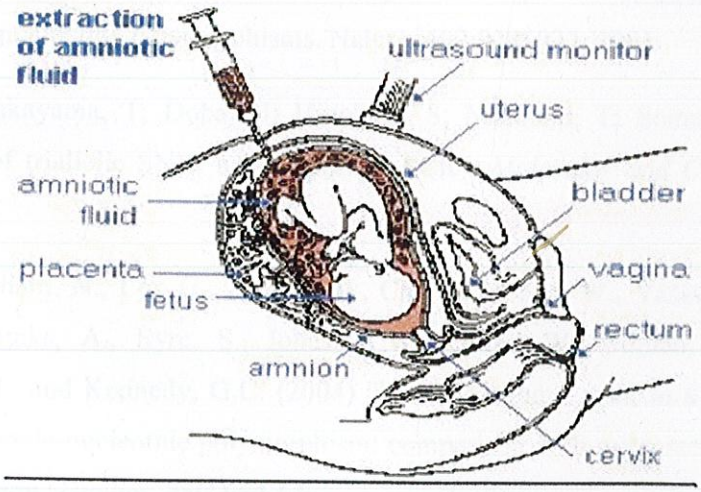


FIG 2.6. Prenatal Diagnosis

2.8 GENETIC CHARACTERIZATION

Genetic marker

A **genetic marker** is a gene or DNA sequence with a known location on a chromosome and associated with a particular gene or trait. It can be described as a variation, which may arise due to mutation or alteration in the genomic loci that can be observed. A genetic marker may be a short DNA sequence, such as a sequence surrounding a single base-pair change -single nucleotide polymorphism, SNP, or a long one, like minisatellites

Work done till date relating to SNPs:

Yue, P; Moult, J (Mar 2006). "Identification and analysis of deleterious human SNPs". *Journal of molecular biology* **356** (5): 1263–74.

R. et al. Sachidanandam. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature, 409:928–933, 2001

Morita, A; Nakayama, T; Doba, N; Hinohara, S; Mizutani, T; Soma, M (june 2007). "Genotyping of triallelic SNPs using TaqMan PCR". *Molecular and Cellular Probes* **21** (3): 171–176.

John, S., Shephard, N., Liu, G., Zeggini, E., Cao, M., Chen, W., Vasavda, N., Mills, T., Barton, A., Hinks, A., Eyre, S., Johes, K.W., Ollier, W., Silman, A., Gibson, N., Worthington, J., and Kennedy, G.C. (2004) "Whole-Genome scan, in a complex disease, using 11,245 single-nucleotide polymorphism: comparison with microsatellites." *American Journal of Human Genetics*. 75(1):54-64.

CHAPTER 3
MATERIALS AND METHODS

3.1 SAMPLE COLLECTION

TABLE 3.1 Blood Samples of the patients suffering from beta-thalassemia

SERIAL NO.	NAME :	PLACE :
1.	Siddharth Gautam	Varanasi
2.	-	Dehradun
3.	Kapil	Dehradun
4.	Mrs. Manjeet	Dehradun
5.	Laxmani	Dehradun
6.	Rahul	Dehradun
7.	Ashish Sharma	Varanasi
8.	Priyam	Varanasi
9.	Nikki Pathak	Varanasi
10.	Amrit	Varanasi
11.	Rohit	Varanasi
12.	Ritu	Varanasi
13.	-	Dehradun
14.	-	Dehradun
15.	-	Dehradun
16.	-	Dehradun
17.	Nikita	Varanasi
18.	Anurag	Varanasi
19.	-	Varanasi
20.	Apurv Sinha	Varanasi
21.	Mehak	Varanasi
22.	-	Varanasi
23.	Monam	Varanasi
24.	-	Dehradun
25.	-	Dehradun
26.	-	Dehradun
27.	-	Dehradun

3.2 GENOMIC DNA ISOLATION

About 27 blood samples were procured from Dehradun and Varanasi. The standard DNA isolation protocol was followed.

3.2.1 MATERIALS REQUIRED

i. Buffer A (100 ml) autoclave

TABLE 3.2. Buffer A

NaCl 100 mM	10 ml
TrisCl (pH 8.0) 1M	5 ml
EDTA (pH 8.0) 1 M	1 ml

ii. Buffer B (100 ml) do not autoclave

TABLE 3.3. Buffer B

NaCl 100 mM	10 ml
TrisCl (pH 8.0) 1M	5 ml
EDTA (pH 8.0) 1M	1 ml
SDS	3 gm

iii. ACD solution (100 ml)

TABLE 3.4. ACD Solution

Citric Acid	0.48 gm
Sodium Citrate	1.32 gm
Glucose	1.47 gm

Dissolve in Milli Q water and filter sterilize (0.25 - 0.50 micron)

iv. Solution I

Phenol (Tris Equilibrated) - Chloroform - Isoamyl alcohol (25 :24 :1)

3.2.2 ***PROTOCOL:*** **Blood DNA isolation**

i. Take 500 µl of blood sample (containing 20% ACD) in sterile 2 ml eppendorf tube

ii. Add 250 µl of buffer A and buffer B each and incubate at 37 °C for 1 -½ hour.

iii. Add Proteinase K to a final concentration of 100 µg/ml.Mix properly by tapping and place at 50°C for another 45 mins. Intermittent mixing is preferred.

iv. Add 1 ml of Solution I. Mix properly by inverting the tube and spin at 10,000rpm for 10 mins at 4 °C.

v. Carefully transfer the aqueous phase to a fresh eppendorf.

vi. If the solution shows turbidity then repeat the above step once or twice depending upon clarity of the solution.

vii. Add about 900 µl of ethanol. Mix and keep at -20° C for overnight precipitation.

viii. Centrifuge at 10,000rpm for 10 mins. Retain the pellet.

ix. Dislodge the pellet using 600 µl of 70% ethanol and centrifuge for 5 mins at room temperature.

x. Dry the pellet and resuspend in 50 µl TE buffer.

3.3 PRIMER DESIGNING

3.3.1 COMPLETE SEQUENCE:

i. Complete sequence of β globin gene was downloaded from NCBI database.

ii. Finally the sequences were put in the Primer 3 software for primer designing.

iii. Total length of gene sequence was about 2kb.

iv. For sequencing the sequence was broken down into 2 parts, 1kb each.

3.3.2 ß- GLOBIN GENE SEQUENCE DEPICTING SNPS

1)-28 A-G

.GCAGGAGCCAGGGCTGGGCATAGAAGTCAGGGCAGAGCCATCTATT

GCTT

2) -29 A-G

.GCAGGAGCCAGGGCTGGGCATGAAAGTCAGGGCAGAGCCATCTATT

GCTT

3) Codon 0 T-G

...GCAGGAGCCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGC

TTACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACC

GTGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGC

AAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAG

4)IVSII 65 4, C-T

ATGTGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATTTTAAAA
AATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTATTTCTAATACTTTCCC
TAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGCCTCTTTGCAC

CATTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGG**T** AATAGCAATATC

TCTGCATATAAATATTTCTGCATATAAATTGTAACTGATGTAAGAGGTTTCAT
ATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGTTG
GGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAAT

5) Codon 43 G-T

.GCAGGAGCCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGC
TTACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCA
TGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAA
GGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTTGGTATCAAGG
TTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGA
GAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTCTATTT
TCCCACCCTTAGG CTG CTG GTG GTC TAC CCT TGG ACC CAG AGG TTC TTT

T AG TCC TTT GGG GAT CTG

3.4 POLYMERASE CHAIN REACTION

TABLE 3.5. PCR Master Mix

MATERIALS	1X
Template DNA (50 ng)	0.5 µl
Primers	0.5 µl each
10X PCR Buffer(with $Mgcl_2$)	2.4 µl
10mM DNTPs	0.3 µl
Autoclaved Water	10.5 µl
Taq DNA Polymerase	0.23 µl
TOTAL	15 µl

3.4.1 PCR PROTOCOL

i. To perform several parallel reactions, the preparation of a master mix containing Water, buffer, dNTPs, primers and *Taq* DNA polymerase in a single tube, which can then be aliquoted into individual tubes. Template DNA solutions is then added.

ii. Gently vortex and briefly centrifuge all solutions after thawing.

iii. Add, in a thin-walled PCR tube, on ice :

iv. Gently vortex the sample and briefly centrifuge to collect all drops from walls of tube.

v. Place samples in a thermocycler and start PCR

vi. Resolve products using 1 X TAE, 2% Agarose gel.

3.4.2 *PCR CYCLE*

TABLE 3.6. PCR Cycle

SERIAL NO.	STAGE	STEP	TEMPERATURE	TIME
1	STAGE 1 *1 CYCLES		95° C	4 mins
2	STAGE 2 *30 CYCLES	STEP 1	95° C	30 sec
		STEP 2	GRADIENT	30 sec
		STEP 3	72° C	2 mins
3	STAGE 3 *1 CYCLES		72° C	10 mins
4	STAGE 4		4° C	∞

3.4.3 *AGAROSE GEL ELECTROPHORESIS*

Genomic DNA was resolved using 0.8% agarose gel and for resolving PCR products 2% agarose gel was prepared.

3.4.4 MATERIALS REQUIRED

i. TAE buffer:

-50X Stock solution of TAE was prepared by adding the following:

TABLE 3.7. TAE Buffer

Tris base	60.5g
Glacial acetic acid	14.275mL
EDTA(0.5M, pH 8.0)	25mL
Distilled water	210mL

Working formula:

$M_1V_1 = M_2V_2$

Where, M = Molarity; V = volume

I 1X TAE was prepared from the 50 X stock solution by substituting the appropriate values in the formula.

II Ethidium bromide (EtBr).

III Gel loading dye : (6 X)- The dye consists of Bromophenol blue, xylenecyanol and glycerol.

IV Agarose salt.

3.4.5 PROTOCOL

I 0.8g of agarose for DNA and 2 g for PCR was taken in a flask and 100 mL of 1X TAE was added.

II The mixture was heated in the microwave for a few minutes.

III The gel was cooled to about 60°C following which 2.5 µL of EtBr was added.

IV The gel was cast in a pre-cleaned gel tray; the comb was set and was allowed to solidify.

I 5 µL of DNA sample was taken on a parafilm and mixed with 3 µL of 6X loading dye

II The combs were gently removed and the DNA was loaded in the wells.

III 100bp DNA ladder was loaded.

IV Electric supply was switched on at 100V.

3.5 ELECTROPHORETIC SEPARATION OF THE PCR PRODUCTS AND PURIFICATION OF AMPLIFIED PRODUCTS

Amplified products from the PCR reactions were purified using Millipore's Montage 96 filter plate as per the manufacturer's protocol (this was performed to remove the unused dNTPs and salts and unused primers), and eluted with 50 µl of sterile nuclease –free dd water following a 30 minute room temperature incubation. Purified PCR products were quantified by A260 readings using a SpectraMax-M2 plate reader (100 µl at 1 :25 dilution). PCR products were stored at -20°C until further sequencing.

3.6 SEQUENCING OF PCR PRODUCTS

The purified PCR products generated from genomic DNA template were sequenced from both ends, using the Forward and Reverse primers using sangers enzymatic method. For the sequencing of beta-globin gene in different genotypes the samples are sent to Xceleris life sciences.

3.7 PROTEIN MODELING

After getting sequencing results of the previously amplified Dehradun Chandigarh samples, Protein Modeling was done to find out the Binding Energy change in the normal and mutated protein of their respective oxygenated and deoxygenated states. The X-ray structure of normal haemoglobin was downloaded from the protein data base and used as template structure to built the mutated haemoglobin structure. The point mutations were introduced in the structure by substituting the mutated amino acid in the pdb file

using molecular builder (Maestro, Schrödinger Inc.). The protein structures were prepared and energy minimized using Protein preparation wizard.

3.7.1 Computational Details

Density functional calculations with the BP86 and B3LYP functional were performed with the Jaguar program. Two different basic sets were used. Most calculations were carried out with basis set I, which corresponds to that labeled as LACVP ** in the program. This implies an effective core potential replacing the 10 innermost electrons of iron. The basis set is valence double- ζ for all atoms and has a polarization shell for all atoms different from iron. An additional set of calculations was carried out with the larger basis set II, which is that labeled as LACV3P**++ in the program. This is a valence triple-ζ basis set. It includes diffuse functions for all atoms and it retains the same polarization shells of basis set I. for the oxygenated systems, basis set I contains 1342 basis functions and basis set II, 2052.

Geometry optimizations were carried out with the BP86 functional and basis set I(BP86/I description). Most of the discussion on energetic is nevertheless based on single-point calculations with the B3LYP functional on the optimized BP86 geometries. The problem of the absolute energetics of the oxygen fixation is a difficult one from a computational viewpoint because of the involvement of different spin states. The energy comparison between different spin states is a challenge for DFT methods, and B3LYP calculations on BP86 geometries seem to provide a reasonable balance. As will be shown below, the BP86 results suggest larger binding energies, but the trends between the different systems considered, which the key issue in this article is, are practically unchanged. The calculations used the unrestricted formalism for the quintet deoxygenated state, with the same ground state described in previous work. Spin contamination was always small. For the singlet oxygenated state, a restricted formalism was used.

Most of the energies presented are potential energies in the gas phase. An additional set of calculations was carried out introducing environmental effects through continuum calculations with Jaguar's Poisson-Boltzmann solver. The dielectric constant is set to 4.0 and the probe radius to 1.40. Zero point energies and entropic

contributions have not been evaluated. The second derivative calculations would be very time consuming, and we think it is reasonable to assume that the additional contribution would be similar for all computed systems. In almost all cases, geometry optimizations were complete except for the frozen chain atoms. In the calculation of the subunit β in the T state, the dihedral angle of the acidic hydrogen of one of the propionic substituent had to be frozen to avoid the formation of a spurious hydrogen bond with the distal histidine.

3.7.2 Computational model

The model system used in these calculations consists of full heme group with all its substituent's, and the four closest amino acids to the iron atom, His E7, Val E11, Phe CD1 for the distal side, and His F8 for the proximal side. The latter is bound to the iron through it's ε nitrogen, and is trans to the coordination site of dioxygen. Its introduction is mandatory in any simple representation of the active center of hemoglobin, and it has been so considered in most previous calculations on the system. The other three amino acids are in the so called distal side of the heme, where dioxygen binds to iron. The importance of this distal side of hemoglobin in dioxygen fixation is well established from experimental studies on mutant systems, and from molecular dynamics calculations. The three amino acids that have been considered are the same in both the α and β subunits. They are sensitive to oxygen fixation because they change notably their positions in the X-ray structure of T and R forms of hemoglobin. Furthermore, it is well known that the distal histidine E7 is able to form a hydrogen bond with dioxygen especially in the α chain.

The structural features of different subunits and states of hemoglobin are introduced in the calculations by constraining the coordinates of selected atoms to the value provided in the Protein Data Bank (PDB). The structures taken from the PDB are that with code 2HHB for the deoxygenated T form and the labeled as 1HHO for the oxygenated R form. The computational approach consisted of freezing the α carbons of the amino acids to the position they have in the PDB structure. The peptidic bonds connecting these α carbons to the rest of protein chains have been suppressed in the calculations, and their valence shell has been saturated with two hydrogen atoms,

which have also been kept at same orientation the C-C and the C-N bonds have in the PDB structure.

The three C-H bond distances were frozen to 1.090 Å. There are thus four frozen atoms for each amino acid, as well as the full heme group (except for the dihedral angle of the propionate mentioned above) and the dioxygen, when present, are completely optimized without any additional constraints in the calculation. The propionate groups attached to the heme have been protonated to keep them in neutral form. We could not introduced the amino acids neutralizing them in our calculations for a reason of size and we assume that this aspect will not be critical for oxygen affinity.

Available X-ray PDB structures for the two possible states, R and T, and the two different subunits, α and β were considered. This gives four different systems was evaluated by introducing additional calculations of "non-PDB" systems where dioxygen was added for T states, or removed for R staets. The oxygen binding energy was estimated by using the simple equation:

$$\Delta E = E_{hexa} - (E_{penta} + E_{O2})$$

The geometry optimization were carried out on a total of eight structures: RαD, RαO, TαD,TαO, RβD, TβD and TβO. An additional pair of the calculations, in the presence and absence of dioxygen, were carried out on a simplified system with no distal amino acids, just an imidazole as the proximal ligand and the substituents of the porphyrin replaced by hydrogen atoms.

STEPS INVOLVED IN PROTEIN MODELING

Open Schrodinger/Maestro software

⇓

Import Structure of oxygenated and deoxygenated β- globin chain

⇓

Go to protein preparation wizard

⇓

Pre-process

⇓

Delete A chain and water

⇓

Optimize

⇓

Minimize

⇓

Open Jaguar

⇓

Optimization

⇓

Give job name

⇓

START

CHAPTER 4

RESULTS

4.1 DNA ISOLATION

Samples were collected from hospitals of Dehradun and Varanasi. Standard DNA isolation protocol was followed. Further modifications were done for better results. At the end 27 DNA samples had on average a concentration of 50-100ng/ul, which was used for further analysis (Fig. 4.1).

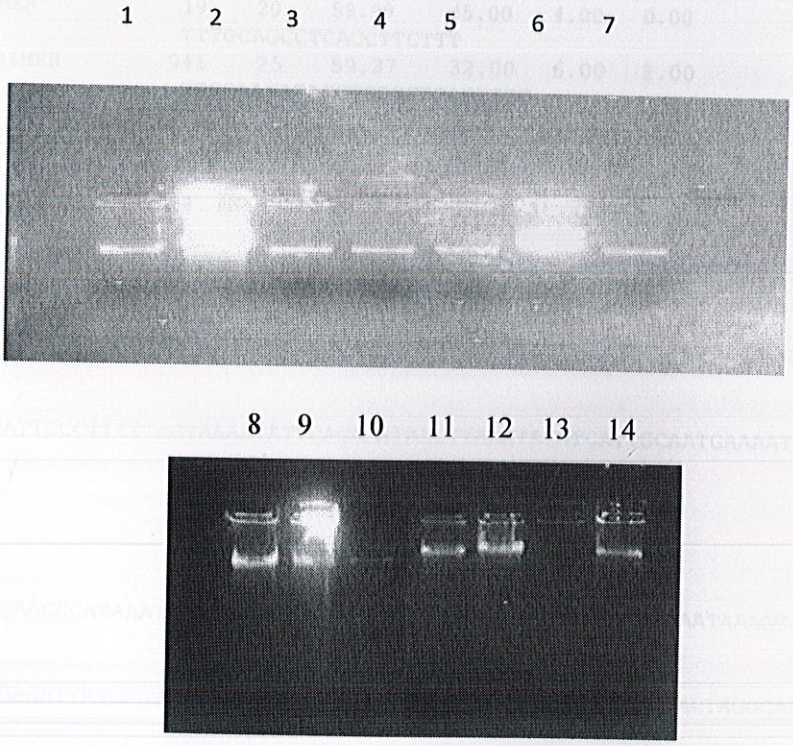


FIG 4.1. DNA Isolation from diseased blood samples

4.2 PRIMERS:

Primers were designed using primer 3. First step in the process of primer designing was breaking the total length of beta-globin genes in two parts, for first primer the a total of 1024 bp length sequence was used and for second primer overlapping 984 bp length sequence was used for primer designing.

4.2.1 PRIMERS FOR THE SEQUENCE 1-1040 bp

```
OLIGO Error! Hyperlink reference not valid.    Error! Hyperlink reference not valid.
Error! Hyperlink reference not valid.  Error! Hyperlink reference not valid.
LEFT  PRIMER          19    20    59.99    45.00    4.00    0.00
                      TTTGCAGCCTCACCTTCTTT
RIGHT  PRIMER        941    25    59.37    32.00    6.00    2.00
                      TTGGAATATATGTGTGCTTATTTGC
SEQUENCE  SIZE:  1040
INCLUDED  REGION  SIZE:  1040

PRODUCT  SIZE:  923,  PAIR  ANY  COMPL:  3.00,  PAIR  3'  COMPL:  0.00

    1  GCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGT

   61  ATTTTCCCAAGGTTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCC

  121  ACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATA



  841  TTTAACCCATAAATATGTATAATGATTATGTATCAATTAAAAATAAAAGAAAATAAAGTA

  901  GGGAGATTATGAATATGCAAATAAGCACACATATATTCCAAATAGTAATGTACTAGGCAG

  961  ACTGTGTAAAGTTTTTTTTTAAGTTACTTAATGTATCTCAGAGATATTTCCTTTTGTTAT

 1021  ACACAATGTTAAGGCATTAA
```

4.2.2 PRIMERS FOR THE SEQUENCE 800-2024 bp

```
OLIGO Error! Hyperlink reference not valid.    Error! Hyperlink reference not valid.
Error! Hyperlink reference not valid.  Error! Hyperlink reference not valid.
LEFT  PRIMER          117    25    59.37    32.00    6.00    2.00
GCAAATAAGCACACATATATTCCAA
RIGHT  PRIMER        1137    20    60.01    55.00    3.00    1.00
ACTCCTAAGCCAGTGCCAGA
SEQUENCE  SIZE:  1224
INCLUDED  REGION  SIZE:  1224

PRODUCT  SIZE:  1021,  PAIR  ANY  COMPL:  4.00,  PAIR  3'  COMPL:  0.00

    1  TGATTTGGTCAATATGTGTACACATATTAAAACATTACACTTTAACCCATAAATATGTAT

   61  AATGATTATGTATCAATTAAAAATAAAAGAAAATAAAGTAGGGAGATTATGAATATGCAA
                                                                >>>>

  121  ATAAGCACACATATATTCCAAATAGTAATGTACTAGGCAGACTGTGTAAAGTTTTTTTTT

  181  AAGTTACTTAATGTATCTCAGAGATATTTCCTTTTGTTATACACAATGTTAAGGCATTAA



1021  CTGCCCTCCCTGCTCCTGGGAGTAGATTGGCCAACCCTAGGGTGTGGCTCCACAGGGTGA

 1081  GGTCTAAGTGATGACAGCCGTACCTGTCCTTGGCTCTTCTGGCACTGGCTTAGGAGTTGG

 1141  ACTTCAAACCCTCAGCCCTCCCTCTAAGATATATCTCTTGGCCCCATACCATCAGTACAA

 1201  ATTGCTACTAAAAACATCCTCCTT
```

4.3 ANNEALING TEMPERATURE DETERMINATION

To amplify the specific DNA fragments from whole genome sequence tests were carried out to determine the exact annealing temperature to increase specificity (Fig. 4.2).

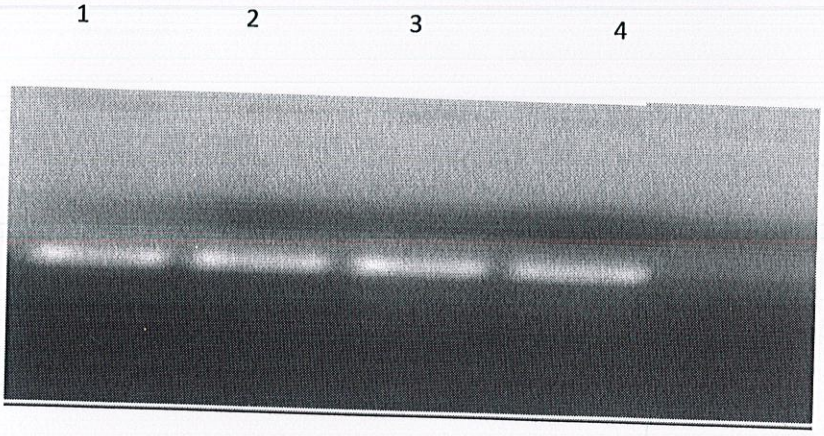


FIG 4.2 Annealing temperature determination for 5th primer (where 1= 56.8^{0}C, 2=57^{0}C, 3=59.6^{0}C and 4= 60^{0}C)

1.4 AMPLIFICATION OF DNA

Amplification of specific DNA fragments from whole genome sequence was carried out at annealing temperature determined with specific primers (Fig. 4.3 ; Fig. 4.4)

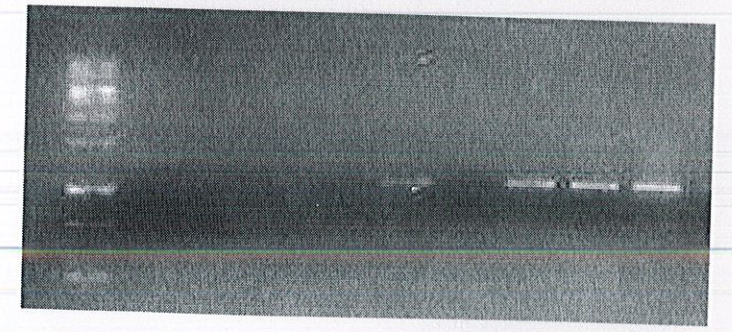


FIG 4.3. Amplification of DNA with different primers (P2); M-100bp ladder

4.5 SEQUENCING RESULTS

We have sent our samples for sequencing after isolating Whole Blood Genomic DNA and And completing its PCR (for amplification). The sequencing results are still awaited.

4.5.1 NOVEL SNPs IN THE 1 SAMPLES

The Different samples showing new mutations. The able shows the samples which exhibit polymorphism i.e the new nucleotide which has replaced the old one and further there location on the gene.

TABLE 4.1 SNPs corresponding to sequences for primer 1

SAMPLE	SNPs POSITION
1. Sample 5	No SNP were detected
2. Sample 9	G-A, T-A, A-T, A-G, T-A, 79,157,176,201,208,214,242, G-A, C-_-, T-G, C-G, T-G 250,254, 280

TABLE 4.2 SNPs corresponding to sequences for primer 2

SAMPLE	SNPs POSITION
Sample 5	No SNP were detected
Sample 9	T-C,A-G,A-N 1255,1450,1688

N -Any Nucleotide not determined yet

Sample 5

```
1 VHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLST      50

1 VHLTPEEKSAVTALWGKVDVDEVGGEALGRLLVVYPWTQRFFESFGDLST      50

51 PDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDP     100

51 PDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDP     100

101 ENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH     146

101 ENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH     146
```

Sample 9

```
1 VHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLST      50

1 VHLTPEGNFAFTALWGNFNVDENGGEALGRLLVVYPWTQRFFESFGDLST      50

51 PDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDP     100

51 PDAVMGNPKVMPLCKKVLGAFSDGLAHLDNLKGSSTILSELHCDKLHVDP     100

101 ENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH     146

101 ENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH     146
```

Fig4.4: Protein sequence alignment between normal and mutated β globin gene

Table4.3: List of previously reported and novel SNPs

SAMPLE 5

POSITION	NORMAL	DISEASED
19	N	D

SAMPLE 9

POSITION	NORMAL	DISEASED
7	E	G
18	V	F
23	V	N
62	A	P
84	T	S
87	T	I

In sample 5; position of mutation in sequence is found 19 and the SNP changes from normal to diseased by change in Asparagine to Aspartic Acid. In sample 9; positions of mutation in sequence that are found are 7, 18, 23, 62, 84, and 87 and the SNPs changes from normal to diseased by Glutamic Acid to Glycine, Valine to Phenylalanine, Valine to Asparagine, Alanine to Proline, Threonine to Serine and Threonine to Isoleucine respectively.

FORMULA OF BINDING AFFINITY:

$$\Delta E = E_{oxy} - (E_{deoxy} + E_{O_2})$$

Table 4.4:Binding energy calculation:

Sample No.		Oxygenated	Deoxygenated	Binding Energy
5	Normal	-3039	-2794	-195
5	Diseased	-2862	-2634	-178
9	Normal	-7529	-6945	-534
9	Diseased	-8436	-7868	-518

In sample no. 5 we were not able to find any novel SNP so the above calculated binding energy are for already reported SNPs and in which the oxygenated binding affinity is -3039, for deoxygenated is -2794 and the binding energy calculated is -195 and this is for normal person but for mutated/affected patient binding affinity for oxygenated is -2862, for deoxygenated is -2634 and its calculated binding energy is -178.

Like wise in sample no. 9 also we few reported as well as novel SNPs by which above calculated binding enrgy is found and in which the oxygenated binding affinity is -7529, for deoxygenated is -6945 and the binding energy calculated is -534; and this is for normal person but for mutated/diseased patient binding affinity for oxygenated is -8436, for deoxygenated is -7868 and its calculated binding enrgy is -518.

CHAPTER5
DISCUSSION

In the genes of thalassemia, there are **15** or more mutations for β-thalassemia. Thirteen of these mutations create new restriction sites of various restriction enzymes. Traditionally, detection of these mutations is dependent on allele-specific hybridization or direct sequencing. In this project, we devised a method by using selective amplification of a DNA fragment of β globin gene with specific oligonucleotide primers, followed by digestion of the amplified product by restriction enzymes that recognize artificially created or naturally occurring restriction sites. In comparison with previous Studies ,we found this method provides a rapid and simple procedure for identifying all the mutations found in the genes. The method is not only very useful for detecting mutations of β - thalassemia and diagnosis , but also for further investigation of new mutations.

DNA was isolated from the samples with the help of the standard protocol. Primers were designed for the whole β globin gene. Primers were also designed for DNA fragments of mutated β globin gene. Annealing temperatures were determined experimentally by PCR for respective primers.

Following are the annealing temperatures for primers:-

1. PRIMER 1	-	55.5 ° C
2. PRIMER 2	-	60 ° C
3. IVSII nt 654	-	56 ° C
4. Codon 0	-	57.6 ° C
5. Codon 43	-	56 ° C
6. Pri (-28)	-	58 ° C

Using these primers DNA fragments for 27 samples were amplified through PCR.
Samples amplified with the help of Primer 1 and Primer 2 were sent for sequencing.

The discovery of large numbers of single nucleotide polymorphisms (SNPs) in genome - scale sequencing initiatives opens new avenues to the study of the genome -wide distribution of diversity and gene based study and its significance. Generally SNPs are highly abundant but their density differs substantially in different regions of a genome and from genome to genome in humans and may be effectively used as markers (Raymond *et al.*, 1999).

Beside doing the amplification of above mentioned samples, work was also done on the previously found SNPs from Dehradun region to find out their oxygen affinities. The oxygen affinity has been evaluated by computing the energy difference between optimized structure of oxygenate and deoxygenated forms of normal as well as mutated protein. For doing this, the Protein Modelling of the mutated sequence and the the binding energy calculations were performed using JAGUAR Program. The calculations in this case have been carried out with basis set I which corresponds to LACVP program. The structural feature of β subunit and oxygenated and deoxygenated states of hemoglobin has been introduced in the calculations by constraining the coordinates of selected atoms to the value provided in PDB.

The oxygen binding energy has been calculated by:-

$$\Delta E = (E_{oxy} - (E_{deoxy} + E_{O2}))$$

Therefore, observing the change in binding energy of both the mutated and normal proteins we found that:-

- The oxygen affinity of beta subunit is different in case of oxygenated and deoxygenated states.
- The binding energy of oxygen is affected because of SNPs in beta globin gene.

CHAPTER 6
CONCLUSION AND FUTURE PROSPECTS

Early comprehensive treatment has changed thalassemia from a fatal pediatric disease to one in which patients live productive lives throughout adulthood. Advances in treatment are exciting, resulting in the potential for cure and improved quality of life. Doctors diagnose thalassemia using blood tests, including a complete blood count (CBC) and special hemoglobin tests. Tests on the amount of iron in the blood to find out whether the anemia is due to iron deficiency or thalassemia.

The present study was aimed to identify *de novo* Single Nucleotide polymorphisms (SNP) in the beta-globin genes. The present study was aimed to identify *de novo* Single Nucleotide polymorphisms (SNP) in the beta-globin genes and to find out the effect of previously found SNPs on the oxygen binding affinity of hemoglobin molecule. The amplified product has been sent for sequencing and the oxygen affinity has been evaluated by computing the energy difference between optimized structure of oxygenate and deoxygenated forms of normal as well as mutated protein. Therefore, the change in binding energies clarifies that Oxygen binding affinity of hemoglobin molecule is also affected due to the presence of SNPs in thalassemic patients. Common and novel SNPs may then directly be taken up for the thalassemia diagnosis and to understand its effect on the severity of the disease.

BIBLIOGRAPHY

Ambekar SS, Phadke MA, Balpande DN, Mokashi GD, Khedkar VA, Bankar MP, Gambhir PS, Bulakh PM, BasutkarDG : The Prevalence and Heterogeneity of Beta Thalassemia Mutations in The Western Maharashtra Population IJHG 1(3): 219 -223 (2001)

Atweh GF, Forget BG: Identification of a beta-thalassemia mutation associated with a novel haplotype of RFLPs. Am J Hum Genet. 38(6): 855-859 (1986).

Christina Vrettou, Joanne Traeger-Synodinos, Maria Tzetis, George Malamis, and Emmanuel Kanavakis1 : Rapid Screening of Multiple Globin Gene mutations by RealTime PCR on the LightCycler: Application to Carrier Screening and Prenatal Diagnosis of Thalassemia Syndromes Clinical Chemistry 49 :5 769-776 (2003)

Ewing, B., Hillier L Wendl M and Green P., Basecalling of automated sequencer traces using Phred I Accuracy assessment Genome Research, 8 (3):175-85 (1998).

Garewal G, Das R, Ahluwalia J, and Marwaha RK : Prevalence of the H63D mutation of the HFE in north India : Apr;74(4):333-6 (2005)

Jan-Gowth Chang, Pao-Huei Chen, Shyh-Shin Chiou, Long-Shyong Lee, Liuh-l Perng, and Ta-Chih Liu: Rapid Diagnosis of β-Thalassemia Mutations in Chinese by Naturally and Amplified Created Restriction Sites. 1992 by the American Society of Hematology.

John, S., Shephard, N., Liu, G., Zeggini, E., Cao, M., Chen, W., Vasavda, N., Mills, T., Barton, A., Hinks, A., Eyre, S., Johes, K.W., Ollier, W., Silman, A., Gibson, N., Worthington, J., and Kennedy, G.C. "Whole-Genome scan, in a complex disease, using 11,245 single-nucleotide polymorphism: comparison with microsatellites." *American Journal of Human Genetics*. 75(1):54-64. (2004)

Ko TM, Hsu PM, Hsieh FJ, Lin KH, Lee SC, Lee TY: DNA polymorphism haplotype analysis of beta-thalassemia genes in Taiwan. 88 (2):103-7 (1989)

Leung NT, Lau TK, Chung TKH "Thalassemia screening in pregnancy". Curr Opinion in Ob Gyn 17 : 129-34 (2005).

Miller RD, Duan S : Efficient High-Throughput Resequencing of Genomic DNA.

Genome Research, 13 :717-720 (2003).

Morita, A; Nakayama, T; Doba, N; Hinohara, S; Mizutani, T; Soma, M "Genotyping of triallelic SNPs using TaqMan PCR". *Molecular and Cellular Probes* **21** (3): 171–176. (2007).

Raymond JC, Michael: Genome-wide mapping with biallelic markers in Arabidopsis thaliana. Nature genetics 23 : 203-207 (1999).

Sachidanandam R, A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature, 409:928–933.(2001)

Tan JA, Tay JS, Kham S, Chin TM, Wong HB, Lai PS. Molecular characterization of ß-thalassaemia in Singaporean Chinese : application to prenatal diagnosis. J Paediatr Child Health;29 :461-463 (1993)

Venkatesan R, Sarkar R, Old JM : Beta-Thalassaemia mutations and their linkage to betahaplotypes in Tamil Nadu in southern India.42 (5):251-6 (1992)

Weatherall DJ, Clegg JB: The Thalassemia Syndromes. (ed 3 New York, NY, Blackwell Scientific, 1981)

www.genome_govlearning About Thalassemia.com

www.MedlinePlus Thalassemia.com

www.wikipedia.com

Yue P, Moult, J "Identification and analysis of deleterious human SNPs". *Journal of molecular biology* **356** (5): 1263–74. (2006).

Zeng YT, Huang SZ: Disorders of hemoglobin in China. J Med Genet. 72(1): 245-78. (1987)