

**Study on the frequency of single nucleotide polymorphism
(rs766432) in *BCL11A* gene and its association with
hematological parameter and HbF level among HbE/ β -
thalassemia patients in Bangladeshi population**



Submitted by: Shadman Sakib Barno

ID: 19176008

October 2021

**Biotechnology Program
Department of Mathematics and Natural Sciences
BRAC University
Dhaka, Bangladesh**

**Study on the frequency of single nucleotide polymorphism
(rs766432) in *BCL11A* gene and its association with
hematological parameter and HbF level among HbE/ β -
thalassemia patients in Bangladeshi population**



**A Dissertation Submitted to the Department of Mathematics and
Natural Sciences, BRAC University in Partial Fulfillment of the
Requirement for the Degree of Master of Science in Biotechnology**

Submitted by: Shadman Sakib Barno

ID: 19176008

October 2021

**Biotechnology Program
Department of Mathematics and Natural Sciences
BRAC University
Dhaka, Bangladesh**

Declaration of authenticity

I, the undersigned, Shadman Sakib Barno, declare that this dissertation is my original work made solely to fulfill the purposes and objectives of this study, that it has not been previously submitted to any other university for a higher degree and that any third-party research referenced within has been given proper citations. It is further declared that this work has been carried out under the joint supervision of Dr. Md. Kaiissar Mannoor, Senior Scientist and Head, institute for developing Science and Health initiatives (ideSHi), Mohakhali, Dhaka as External Supervisor and Professor Dr. Mahboob Hossain, PhD, Professor, Department of Mathematics and Natural Sciences, BRAC University, Mohakhali, Dhaka as Internal Supervisor.

Shadman Sakib Barno

Shadman Sakib Barno

Kaiissar Mannoor

Candidate Certified **Dr. Md. Kaiissar Mannoor**
(External Supervisor)
Senior Scientist and Head
Institute for developing Science & Health
initiatives (ideSHi)
IPH Building (2nd Floor)
Mohakhali, Dhaka-1212,
Bangladesh.
E-mail: **kaiissar@ideshi.org**

Mahboob Hossain

Professor Dr. M. Mahboob Hossain, PhD
(Internal Supervisor)
Professor
Department of Mathematics and
Natural Sciences
BRAC University
66, Mohakhali, Dhaka-1212,
Bangladesh.
E-mail: **mmhossain@bracu.ac.bd**

Approval

The thesis titled “Study on the frequency of single nucleotide polymorphism (rs766432) in *BCL11A* gene and its association with hematological parameter and HbF level among HbE/ β -thalassemia patients in Bangladeshi population” submitted by:

Shadman Sakib Barno (19176008)

has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Master of Science in Biotechnology on

Examining Committee:

Supervisor: (Member)	Professor Dr. M. Mahboob Hossain, PhD Professor, Mathematics and Natural Sciences Brac University
Co-Supervisor: (Member)	Dr. Md. Kaiissar Mannoor Senior Scientist and Head Institute for developing Science & Health initiatives (ideSHi)
Program Coordinator: (Member)	Full Name Designation, Department Institution
External Expert Examiner: (Member)	Full Name Designation, Department Institution
Departmental Head: (Chair)	Full Name Designation, Department Institution

Ethics Statement

This material is original work which has not been previously published elsewhere. It contains my own research and analysis done solely for this material. The paper properly credits all sources used with proper citation.

Dedicated to
my family and well-wishers

Acknowledgment

I am grateful for my circumstances in being given the opportunity to complete my MS thesis at the facilities of BRAC University, and at ideSHi (Institute for Developing Science and Health initiatives). The members of both institutions have been very helpful and supportive, and the completion of this thesis would not have been possible without them.

In particular, I am thankful to the following individuals from the Department of Mathematics and Natural Sciences at BRAC University: Professor A F M Yusuf Haider, PhD, Professor and Chairperson of the Department of Mathematics and Natural Sciences, for making such opportunities possible for students, for contributing to the overall quality of scientific education in BRAC University, and for overseeing the smooth function of the department; Assistant Professor Iftexhar Bin Naser, PhD, for overseeing the submission and presentation of theses in the department; and Professor Mahboob Hossain, PhD, Professor of the Department of Mathematics and Natural Sciences for acting as my supervisor within BRAC University and playing a key role in the successful completion of this thesis.

Additionally, I am thankful to the following individuals for their generous assistance during my time spent at ideSHi: Dr. Firdausi Qadri, Chairperson; Prof. Dr. Syed Saleheen Qadri, Scientific Coordinator; Dr. Md. Kaiissar Mannoor, Head & Senior Scientist; Suprovath Kumar Sarker, Research Fellow; Farjana Akther Noor, PhD Fellow; Golam Sarower Bhuiyan, Senior Research Officer; Mohabbat Hossain, Research Officer; Khalid Arafat Aubboy, MS Thesis Student; Subarna Sayed Ety, MS Thesis Student; Marzia Tasnim Mitu, MS Thesis Student; and Jannat Swarna, MS Thesis Student.

The contributions of everyone listed ranged from answering whatever queries I had regarding this thesis, to providing valuable guidance and feedback on it, and to providing equally valuable assistance in my research and data collection. This thesis would have been impossible to complete without all of their help, especially in light of certain world events, as well as unexpected health issues on my part which required an all too long recovery.

Shadman Sakib Barno
October 2021

Abstract

Background: Thalassemia is a genetic disorder that causes malformed globin chains in hemoglobin and is the most prevalent blood disorder worldwide with high incidence rates even in Bangladesh. When β -thalassemia is inherited together with a hemoglobin E allele, the resulting condition HbE/ β -thalassemia is sometimes characterized by a severe, transfusion-dependent thalassemia major. The aim of this study was to find the frequency of single nucleotide polymorphism (SNP) rs766432 in the *BCL11A* gene and its association with hematological parameters and HbF production among patients with HbE/ β -thalassemia in Bangladesh's population.

Methods: A total of 63 HbE/ β -thalassemia patients from the Thalassemia Samity Hospital, Dhaka participated in this study, with the full consent of all parties involved. Complete Blood Count (CBC) and hemoglobin electrophoresis were performed to determine hematological parameters and hemoglobin fractions respectively. SNP rs766432 frequency was determined through Sanger Sequencing followed by High Resolution Melt Curve Analysis based on Real-Time PCR. Graphpad Prism version 7 was used to carry out statistical analysis.

Results: For the HbE/ β -thalassemia patients, the frequency of rs766432 in *BCL11A* was found to be 75% homozygous genotype AA, 25% heterozygous genotype CA. In healthy individuals on the other hand, 74% homozygous major allele AA and 26% heterozygous minor allele CA was observed. However, no statistically significant differences were found in the mean $\pm$ SD of hemoglobin indices for the CA genotype and AA genotypes. In addition, neither the age at first blood transfusion nor the transfusion interval was significantly higher in patients with the rs766432 polymorphism in comparison to those without.

Conclusion: This study found SNP rs766432 to have similar frequency among healthy individuals and HbE/ β -thalassemia patients in Bangladesh. However, rs766432 SNP in *BCL11A* gene showed no correlation with either the HbF levels or clinical presentation of HbE/ β -thalassemia patients in the study population.

Keywords: HbE/ β -Thalassemia; *BCL11A* gene; SNP; Fetal hemoglobin.

Table of Contents

Contents

Declaration of authenticity	i
Approval	i
Ethics Statement	ii
Acknowledgment	i
Abstract	ii
Table of Contents	iii
Contents	iii
List of Tables	vi
List of Figures	vii
INTRODUCTION	1
Background	1
Basic facts about Thalassemia	2
Basic facts about HbE/ β -thalassemia	2
Structure of Human Hemoglobin	3
Epidemiology of HbE/ β -Thalassemia	5
Pathophysiology of HbE/ β -Thalassemia	7
Clinical severity & variability of hemoglobin E/ β -thalassemia	7
Genome-wide association studies on <i>BCL11A</i>	8
Repressor mechanism of <i>BCL11A</i> inducing HbF	10
Mechanism for increased HbF production	11
Therapies for HbF induction	11
Aims & Objectives	12
1.12.1 Aims	12

1.12.2 Specific Objectives	12
MATERIALS AND METHODS	14
Study Design	14
Study site and ethical approval.....	14
Sample collection site	15
Study participants	15
Specimen collection and storage.....	15
Hematological Analysis	15
Complete Blood Count (CBC) analysis	15
Hemoglobin Electrophoresis.....	16
Molecular Analysis	16
DNA Isolation	16
Principle	16
Materials & Equipment.....	16
Reagents.....	17
DNA isolation procedure.....	17
Isolated DNA quantification.....	18
Polymerase chain reaction (PCR).....	19
Principle	19
Equipment & Materials.....	20
Reagents.....	21
Procedure.....	21
Primers	21
PCR Reaction.....	22
Agarose Gel Preparation for Gel Electrophoresis	23
Principle	23
Equipment and reagents.....	24

Equipment.....	24
Reagents.....	24
Preparation of Gel and Gel Run of PCR Products	24
PCR Product Purification.....	25
Principle	25
Equipment and reagents.....	26
Equipment.....	26
Materials	26
Reagents.....	26
Purification procedure	26
Sanger Sequencing.....	27
Principle	27
Procedure.....	28
Sequencing analysis	28
Real-Time PCR-High Resolution Melt (HRM) Curve Analysis	29
Principle	29
Equipment & Reagents	29
Equipment.....	29
Reagents.....	30
Procedure.....	30
Statistical Analysis.....	31
RESULTS	32
Demographic Information	32
Hematological status of the study patients	32
Hemoglobin status of the study patients.....	33
PCR results targeting <i>BCL11A</i> rs766432 among the study patients	33
Frequency of <i>BCL11A</i> rs766432 polymorphism among the study participant	33

Allele frequency among HbE/beta thalassemia patients	34
Allele frequency among healthy Individuals.....	34
Comparative analysis of Hb electrophoretic indices between CA and AA alleles of <i>BCL11A</i> rs766432 among the study patients	35
Correlation of the SNP rs766432 in <i>BCL11A</i> gene with age of first blood transfusion and transfusion intervals	37
DISCUSSION.....	39
CONCLUSION.....	42
REFERENCES.....	43

List of Tables

Table 1. List of primers used for amplification of BCL11A gene.....	22
Table 2. Composition of PCR reagent for amplification of BCL11A gene using rs766432 primers	22
Table 3. Thermal cycle temperatures for HRM	30
Table 4. Demographic status of the HbE- β thalassemia patients.....	32
Table 5. Hematological status of the HbE- β thalassemia patients.....	33
Table 6. Hemoglobin fractions among the study patients.....	33
Table 7. Comparison of Hb variants between two allele BCL11A rs766432 among the study patients	35
Table 8. Correlation of SNP rs766432 of BCL11A gene with age of first transfusion and transfusion interval.....	38

List of Figures

Figure 1.	Structure of Hemoglobin (http://www.google.com/images)	4
Figure 2.	The developmental stage and specific regulation of different types of hemoglobin (http://www.google.com/images)	5
Figure 3.	Global distribution of HbE trait and β -thalassemia (¹ Weatherall and Clegg 2001, ² Vichirsky 2005)	6
Figure 4.	The <i>BCL11A</i> gene on chromosome 2p15. Genomic organization of the <i>BCL11A</i> gene on chromosome 2p15 (Thein et al 2009).....	9
Figure 5.	The <i>BCL11A</i> -HbF repression axis (Bauer et al 2015)	10
Figure 6.	Proposed mechanism for increased HbF production in β -thalassemia syndromes. (Rees et al 1999)	11
Figure 7.	Workflow of the study	14
Figure 8.	Polymerase Chain Reaction	20
Figure 9.	Thermal cycling conditions for PCR of <i>BCL11A</i> (rs76) polymorphic region	23
Figure 10.	PCR products of <i>BCL11A</i> rs766432 are visualized on 1.5% agarose gel..	25
Figure 11.	Sanger Nucleotide Sequencing	28
Figure 12.	Sequence chromatogram showing the SNP C>A results in Homozygous major AA in <i>BCL11A</i> rs766432.....	31
Figure 13.	Allele Distributions of <i>BCL11A</i> rs766432 among HbE/beta thalassemia patients	34
Figure 14.	Allele Distributions among Healthy Individuals of <i>BCL11A</i> rs766432	34
Figure 15.	Comparison of HbA between two allele of rs766432.....	35
Figure 16.	Comparison of HbA2 between two allele of rs766432.....	36
Figure 17.	Comparison of HbF between two allele of rs766432	36
Figure 18.	Comparison of HbE between two allele of rs766432	37
Figure 19.	A Graphical Representation of the Mean $\pm$ SD of (A) Age of first blood transfusion (months) (B) Blood transfusion interval (days) of different rs766432 of <i>BCL11A</i> genotypes	38

CHAPTER-01

INTRODUCTION

Background

Thalassemia (also known as Mediterranean anemia, Cooley's anemia, Beta-thalassemia, or Alpha thalassemia) is an inherited blood disorder characterized by malformation of hemoglobin. This specific type of blood disease results in excessive destruction of red blood cells which in turn leads to anemia (Weatherall et al 1996). The two main types of thalassemia are α and β -thalassemia. Phenotypically there are two forms of β -thalassemia: β^0 with no β -globin chain synthesis, and β^+ with some β -globin chain synthesis. Clinically, thalassemia presents as beta-thalassemia trait or minor (β^+ or β^0), intermedia (β^+/β^+ ; β^+/β^0) or major (β^0/β^0) (George and Khuziah 1994). When β -thalassemia is inherited together with a hemoglobin E (HbE) allele, the resulting condition, HbE- β thalassemia (HbE- β Thal), is sometimes characterized by a severe, transfusion-dependent thalassemia major (Weatherall 2000). HbE/ β -thalassemia accounts for up to half of severe cases of β -thalassemia. HbE/ β -thalassemia is common in India, Bangladesh, and Southeast Asia, but as a result of migration, is not limited to those geographical areas alone (Algiraigri et al 2017). The transcription factor B-cell lymphoma/leukemia 11A (*BCL11A*) is the major fetal hemoglobin (HbF) silencer. Though, the repression mechanism of γ -globin gene expression by *BCL11A* has not been fully unraveled (Yang et al., 2019). SNPs in *BCL11A* were associated with F-cell numbers in Chinese with β -thalassemia trait, and with sickle cell anemia. The functional motifs reside within a 3 kb region in the second intron of *BCL11A* which is responsible for modulating F-cells and HbF levels (Sedgewick et al 2008). A drug named Hydroxyurea can enhance hemoglobin F production. Therefore, disease modification through Hydroxyurea treatment may be possible. Further exploration is needed to determine the long-term effects of hydroxyurea therapy (Fucharoen et al 2000).

Basic facts about Thalassemia

Thalassemia is characterized by the absence or reduced synthesis of globin chains α , β , δ , γ , ζ and ϵ in human hemoglobin (Hb). The two main types of thalassemia are α - and β -thalassemia (George and Khuziah 1994). The highest prevalence of the disease is in the 'thalassemia belt' which includes the Mediterranean region, parts of the Middle East, the Indian subcontinent, the southern parts of the far East, Pakistan, and Southeast Asia (Usman et al 2009). At least 5.34% of the current global population are carriers of a significant hemoglobin variant (Model et al 2008). In 2010, the combined populations of India, Pakistan, Bangladesh, and Sri Lanka totaled 1.56 billion, representing 23% of the global population (PRB 2010).

Thalassemia is the most common hereditary disorder in the world, constituting a major health problem globally. The problem is particularly severe in developing countries, where resources are limited. With child mortality rates from diarrhea, respiratory tract infections and malnutrition declining due to access to better healthcare and nutrition, more children are now surviving long enough to require treatment for hereditary disorders like thalassemia. It has been observed that a need for genetic services arises for a country when infant mortality rates fall below 46/1,000 livebirths. Bangladesh's infant mortality rate is now at 46/1,000 livebirths and falling. Therefore, thalassemia and other genetic disorders will soon become a major public health problem for this country (Khan et al 2005).

The World Health Organization (WHO) estimates that at least 6.5% of the world population are carriers of different inherited disorders of hemoglobin (Modell 1995). Another WHO report estimates that in Bangladesh, 3% are carriers of β -thalassemia and 4% are carriers of HbE / β -thalassemia. More than 7,000 children are born with thalassemia in Bangladesh each year (Khan 2006). Patients suffering from β -thalassemia major and HbE / β -thalassemia do not survive for more than 5 years without blood transfusions (Shekhar et al 2007).

Basic facts about HbE/ β -thalassemia

HbE/ β -thalassemia is among the most common forms of hemoglobinopathies worldwide (Weatherall and Clegg 1996). HbE is a splice variant of normal β -globin

protein. This cryptic splice site (related to HbE) is not normally used for mRNA processing. This new splice site competes with the normal splice site and thus produces a protein where Lys substitutes Glu at position 26 (Orkin 1982). This variant (HbE) is the protein produced as a result of this mutation. HbE can be present in both heterozygous and homozygous states, but it can interact with other β -thalassemia types to produce HbE / β -thalassemia. Several mutations, which interact with mRNA processing, include abnormalities located in the exon-intron junctions affecting donor or receptor sites. When one of the nucleotides from the GT/ . . . /AG junction is modified, the result is a total abolition of the splicing leading to a β^0 -thalassaemic defect. Other mutations have been described within the splicing consensus sequences that seem to play a role in selection of the correct splicing site. HbE is not the mutated protein, but the fact is that its β -subunits are synthesized in too low a quantity to bind all the available alpha chains, resulting in mild β^+ -thalassemia. The primary clinical importance of HbE trait arises when the β^E allele interacts with other β -thalassemia mutations leading to moderate-to-severe anemia known as HbE / β -thalassemia (Ong-Ajyooth et al 1987).

HbE/ β -Thalassemia is thought to have first originated in Southeast Asia (Krishnamurti 2000). In Sri Lanka, approximately 40% of HbE / β -Thalassemia patients are expected to require treatment at any time due to ambiguities resulting from its phenotypic variability (Olivieri et al 2008).

Structure of Human Hemoglobin

Hemoglobin is a 64.4 kD tetramer consisting of two pairs of globin polypeptide chains: one pair of alpha-like chains, and one pair of non-alpha chains (Figure-1). The chains are designated by Greek letters, which are used to describe the particular hemoglobin (e.g., Hb A is α^2/β^2). Two copies of the alpha-globin gene (HBA2, HBA1) are located on chromosome 16 along with the embryonic zeta genes (HBZ). There is no substitute for alpha globin in the formation of any of the normal hemoglobins (Hb) following birth (e.g., Hb A, Hb A2, and Hb F) (Steinberg 2011).

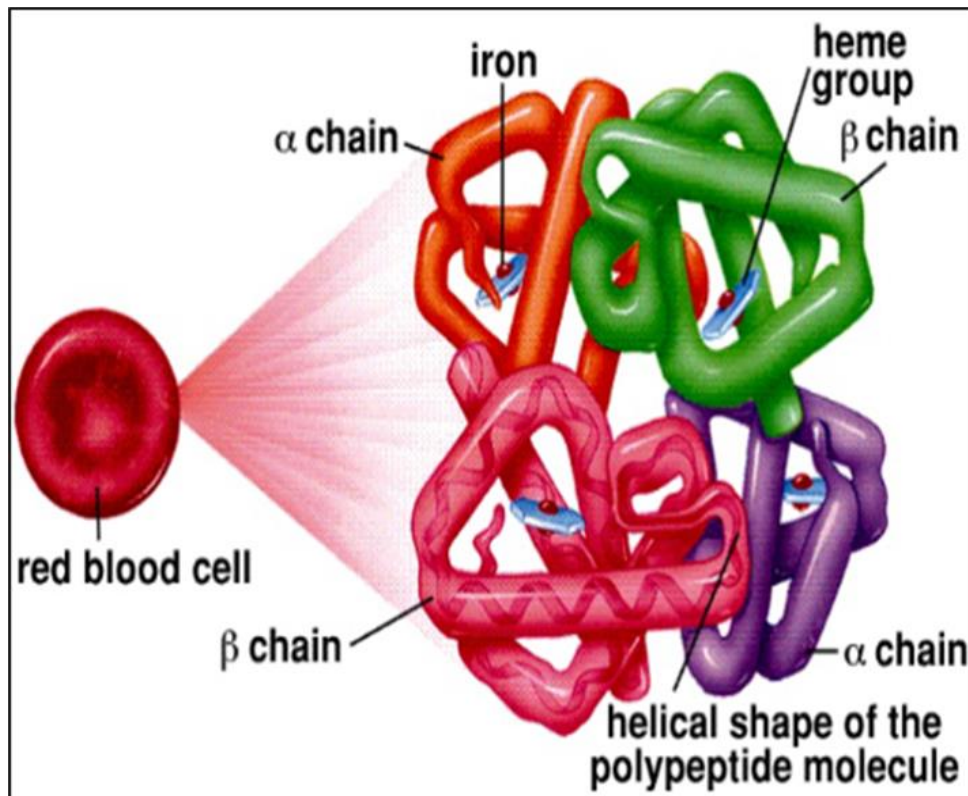


Figure 1. Structure of Hemoglobin (<http://www.google.com/images>)

Embryonic hemoglobins that contain zeta (HBZ)- or epsilon (HBE1)-globin chains are primarily the product of yolk sac erythroblasts and are detectable only during the very earliest stages of embryogenesis. In early development, between weeks 4 and 14 of gestation, the human embryo synthesizes three distinct hemoglobins. These embryonic hemoglobins are in their order of formation (Hofmann et al 1995).

By approximately the 14th week of gestation, after establishment of erythropoiesis in the fetal liver and spleen, these embryonic hemoglobins are replaced by HbF. After the 8th week of gestation, HbF becomes the predominant hemoglobin of the fetus, and its levels continue to increase until midway through gestation. The concentration of HbF increases with decreasing gestational age and is regulated throughout development (Figure-2). The concentration of HbF in an infant born at 28 weeks of gestation is approximately 90% and decreases to approximately 60% at 10 weeks after birth, a value similar to that of a full-term infant born at 38 weeks (Steinberg 2011).

The transition from HbF to HbA appears to be controlled via a multi-protein complex, which includes expression of *BCL11A* (B-cell lymphoma/leukemia 11A), a

gene which encodes a suppressor of HbF expression (Sankaran, V.G. et al, 2008., Sankaran, V.G. et al, 2010). Of interest, hereditary persistence of HbF (HPFH), in which persistently high levels of HbF continue into adulthood, has been attributed in one family to haploinsufficiency for KLF1 (Borg, Papadopoulos et al 2010).

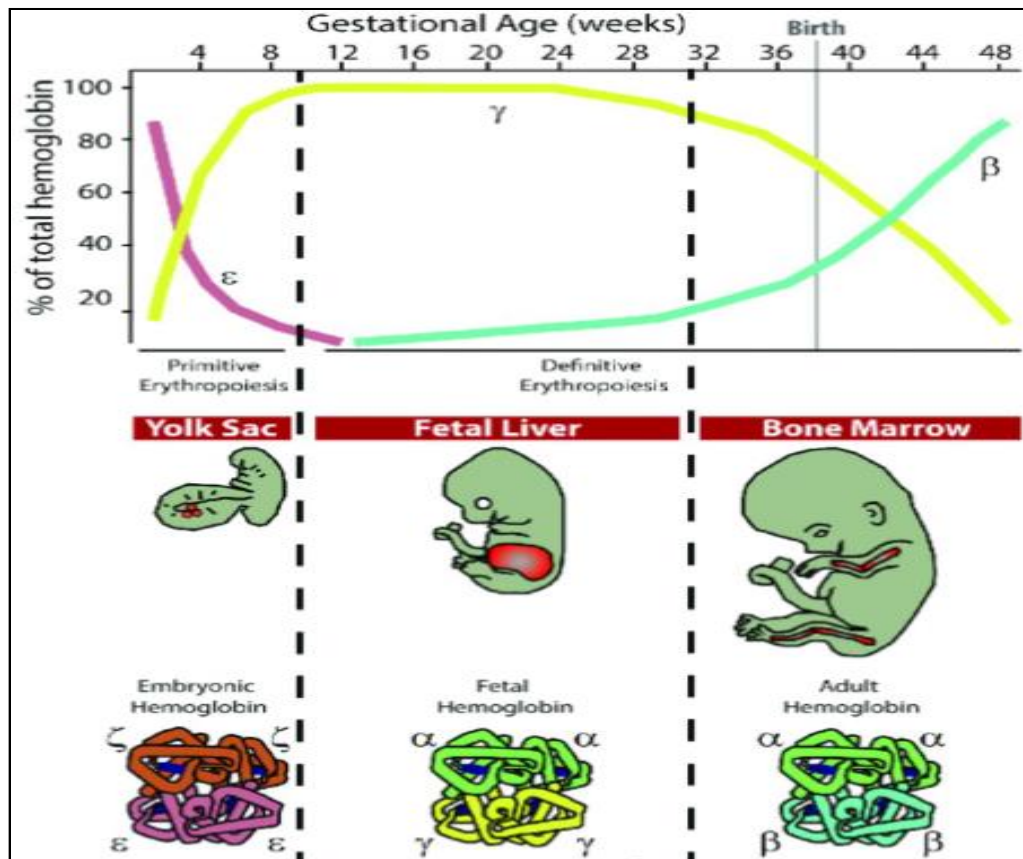


Figure 2. The developmental stage and specific regulation of different types of hemoglobin (<http://www.google.com/images>)

Epidemiology of HbE/β-Thalassemia

Worldwide, patients with hemoglobin E-beta-thalassemia (Hb E/β-thalassemia) represent approximately 50% of those affected with severe beta thalassemia (Modell and Darlinson 2008) (Figure-3). The highest frequencies are observed in India, Bangladesh and throughout Southeast Asia, particularly in Thailand, Laos, and Cambodia, where it is common for individuals to inherit alleles for both haemoglobin E (Hb E) and beta-thalassemia (Weathall and Clegg 2001). Throughout these regions, Hb E/β-thalassemia has become an increasingly severe public health problem. In Thailand, about 3,000 affected children are born annually, and there are an estimated 100,000 patients living with the condition (Flint et al 1998). HbE/β-thalassemia is also common in Indonesia and Sri Lanka, and while previously rarely

diagnosed in North America or Europe, has recently become the most common form of β -thalassemia identified by many newborn screening programs (Nancy et al 2011).

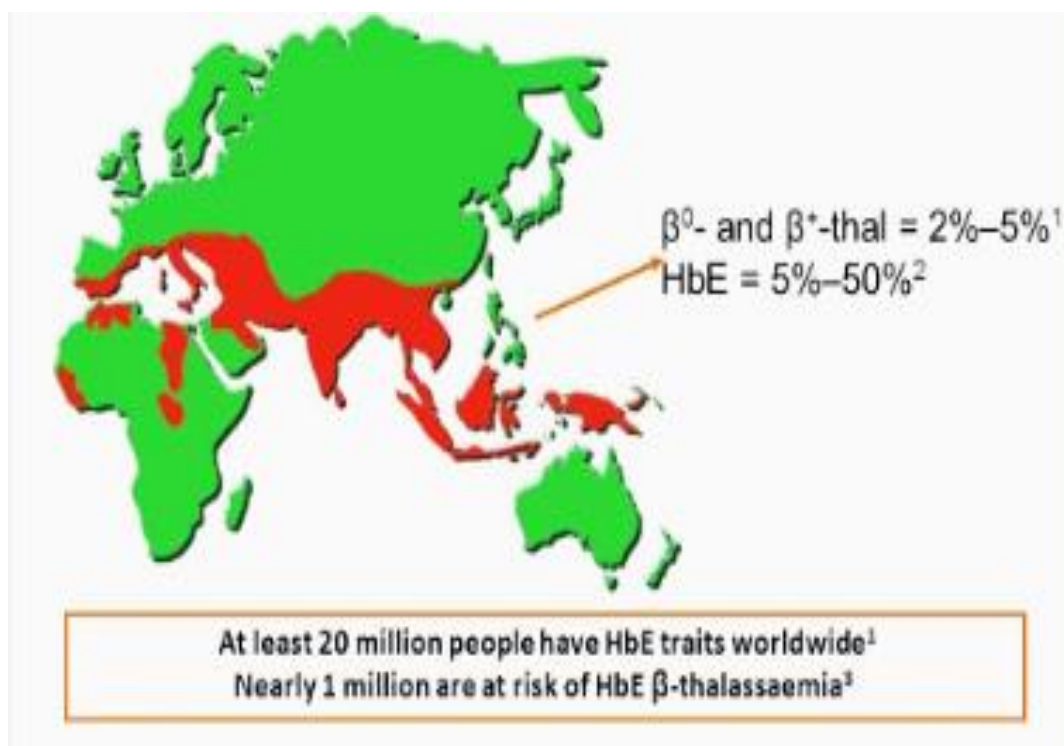


Figure 3. Global distribution of HbE trait and β -thalassemia (¹Weatherall and Clegg 2001, ²Vichirsky 2005)

According to World Health Organization (WHO) estimates, approximately 3% of Bangladesh's population are carriers of beta-thalassemia and 4% are carriers of hemoglobin E (HbE) (WHO, WH et al 2007).

Among tribal children, the prevalence of beta-thalassemia trait was almost identical, but HbE was much higher (41.7%). Another study with a small sample size also observed similar (39 - 47%) prevalence rate of HbE among a tribal population in Bangladesh (Shannon et al 2015). Hence, the prevalence of hemoglobinopathies in Bangladesh could be extrapolated from population-based studies conducted in West Bengal. A recent large population-based study (n=50,487) in rural West Bengal revealed the carrier rate of beta-thalassemia and HbE to be 6.61% and 2.78% respectively (Mandal et al 2014). Another study (n=9,990) estimated the frequency at 3.64% for beta thalassemia traits and 3.92% for the HbE trait (Mohanty et al 2013). The number of individuals with thalassemia (beta major and HbE beta), of varying levels of severity, is estimated to be approximately 60,000 to 70,000 in

Bangladesh. With a birth rate of 21.6/1,000, it could be estimated that nearly 2,500 thalassemia major cases are added every year in Bangladesh (Mandal et al 2014).

Pathophysiology of HbE/ β -Thalassemia

The pathophysiology of HbE/ β -thalassemia is complex with many factors, including reduced β chain synthesis leading to globin chain imbalance (Datta et al 2006). Ineffective erythropoiesis, apoptosis, and oxidative damage are central components of the disease and its shortened red cell survival (Weatherall et al 2001; Datta et al 2006; Pootrakul et al 2000). The instability of HbE is a minor factor in its pathophysiology, but in febrile patients, may account for accelerated hemolysis (Jetsrisuparb et al 2006). HbE/ β -thalassemia results from co-inheritance of a β -thalassemia allele from one parent and the structural variant Hemoglobin E from the other. Hemoglobin E results from a G→A substitution in codon 26 of the β globin gene, which produces a structurally abnormal hemoglobin in addition to activating a cryptic splice site, resulting in abnormal messenger RNA (mRNA) processing. The level of normally spliced mRNA, β E, is reduced (Orkin, S.H. et al, 1982), and, because a new stop codon is generated, the abnormally spliced mRNA is nonfunctional. Hence, hemoglobin E is synthesized at a reduced rate, acting like a mild form of β -thalassemia.

Clinical severity & variability of hemoglobin E/ β -thalassemia

There is an emerging understanding of the genetic and environmental factors that influence the clinical course and severity of anemia in HbE/ β -thalassemia. Genetic factors include the type of β -thalassemia mutation co-inherited with Hb E, the co-inheritance of alpha (α) thalassemia, and polymorphisms shown to be associated with increased synthesis of fetal hemoglobin (Nancy, Zahra and Elliott 2011).

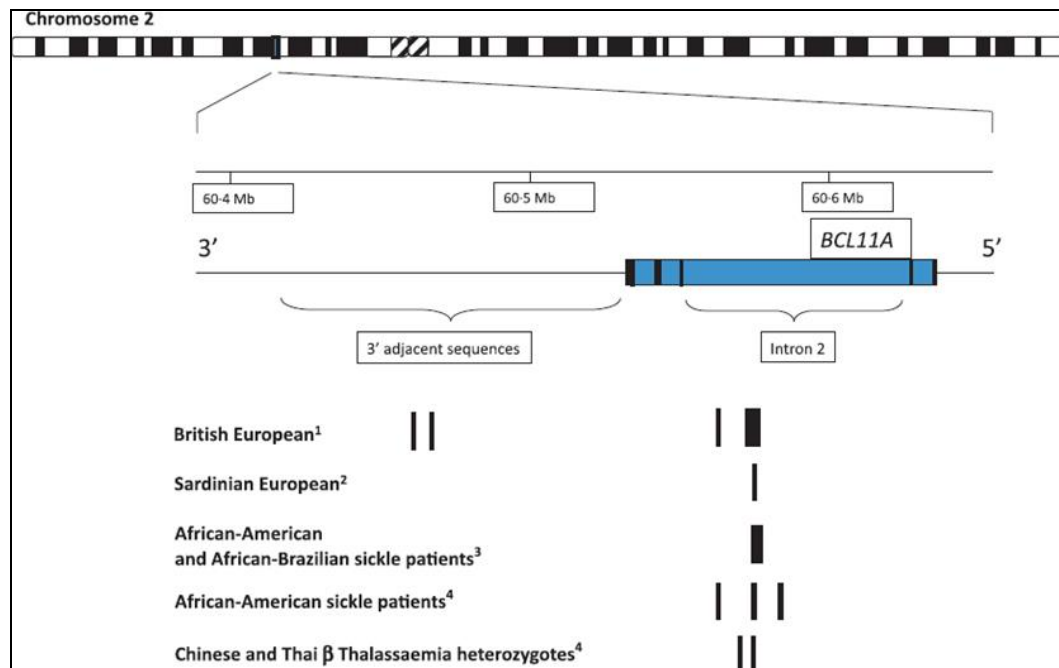
In early studies, Thai investigators (Winichagoon et al 1993) suggested that patients who co-inherit a mild β -thalassemia allele with HbE may have disease on the mild end of the spectrum, while those who co-inherited severe β^+ or β^0 -thalassemia alleles might be more severely affected. However, a less severe β -thalassemia mutation may be only one of several genetic modifiers. Winichagoon et al, defined the mutations in 90 Thai patients with steady-state hemoglobin concentrations

varying from 4.2 to 12.6 g/dl, the same severe β -thalassemia mutation being observed in both mild and severely anemic patients.

Recent studies supporting the beneficial effect of α -thalassemia on HbE/ β -thalassemia also indicate that, as with the beta globin gene mutation, the influence of this factor on the clinical severity of HbE/ β -thalassemia may not be altogether straightforward. Co-inheritance of α -thalassemia appears to be a major genetic factor modifying clinical phenotype, including the finding that a considerable number of patients who co-inherit α -thalassemia may be diagnosed later in life. Clearly, while mild HbE/ β -thalassemia may occur in the absence of α -thalassemia, the presence of α -thalassemia genes should always be considered in mild cases of HbE/ β -thalassemia (Nancy et al 2011).

Genome-wide association studies on *BCL11A*

The chromosome 2p QTL maps to the gene encoding zinc-finger protein *BCL11A*, previously implicated in leukemias and lymphoma pathogenesis (Figure-4). This regulating gene is a major contributor to HbF and F cell variability in Europeans, accounting for 15% of the overall trait variance. The second study involved over 4,000 individuals from the Mediterranean island of Sardinia which showed strong association with HBS1L/MYB and *BCL11A* loci in addition to the HBB locus. Both studies show the striking level of robustness and power that can be reached with new genome-wide association approaches.



Repressor mechanism of *BCL11A* inducing HbF

The transcription factor B-cell lymphoma/leukemia 11A (*BCL11A*) is known as a major HbF silencer (Figure-5). A recent study showed the molecular mechanisms of *BCL11A* binding to -155 HPFH region of γ -globin gene through C2H2 zinc fingers (Yang Y et al., 2019). In erythroid cells, multiprotein transcriptional complexes are participated by *BCL11A* and various protein partners, e.g., DNA-binding erythroid transcription factors (TFs) such as GATA1, FOG1, RUNX1, IKZF1, and SOX6; chromatin regulators and transcriptional corepressors are associated in this regulation, such as, members of the NuRD, LSD1/CoREST, NCoR/SMRT, SIN3, and SWI/SNF complexes. *BCL11A* binds to Locus Control Region (LCR) and also numerous sites at beta-globin gene cluster and promotes interaction between LCR and Beta-globin gene cluster but doesn't bind to β -globin and γ -globin gene clusters themselves. Inhibition of many protein partners in *BCL11A* gene resulting HbF induction (Bauer et al 2015).

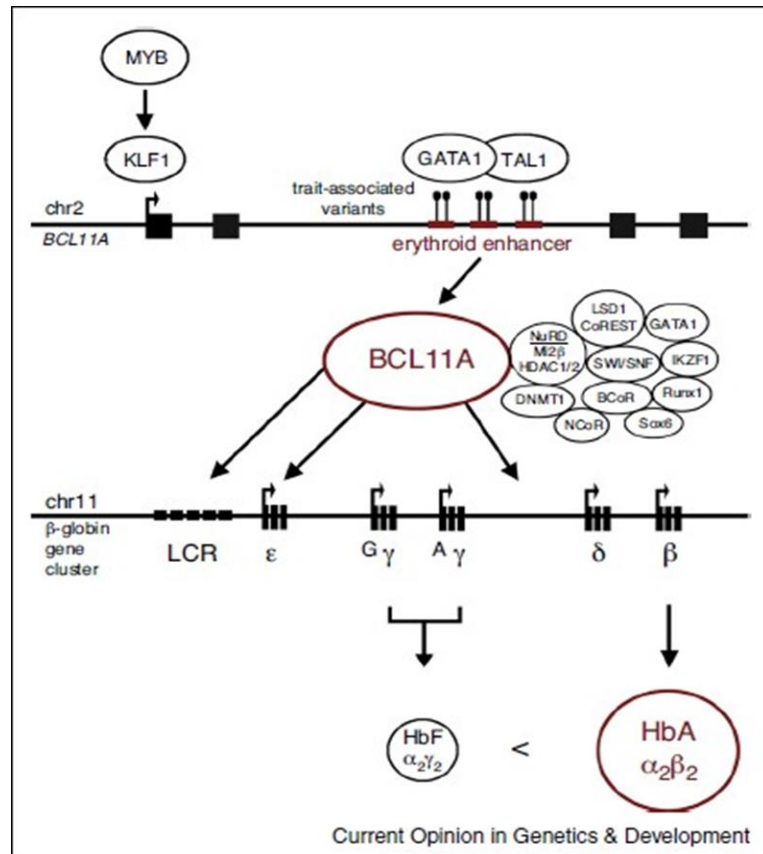


Figure 5. The *BCL11A*-HbF repression axis (Bauer et al 2015)

Mechanism for increased HbF production

According to Rees DC, et al., 1999, the increase in red blood cell precursors is proportional to the expansion of F-cell progenitors. Having less imbalanced globin chains, F-cell precursors have selective advantages which can be observed in increased production of HbF. The mechanism is summarized below (Figure-6):

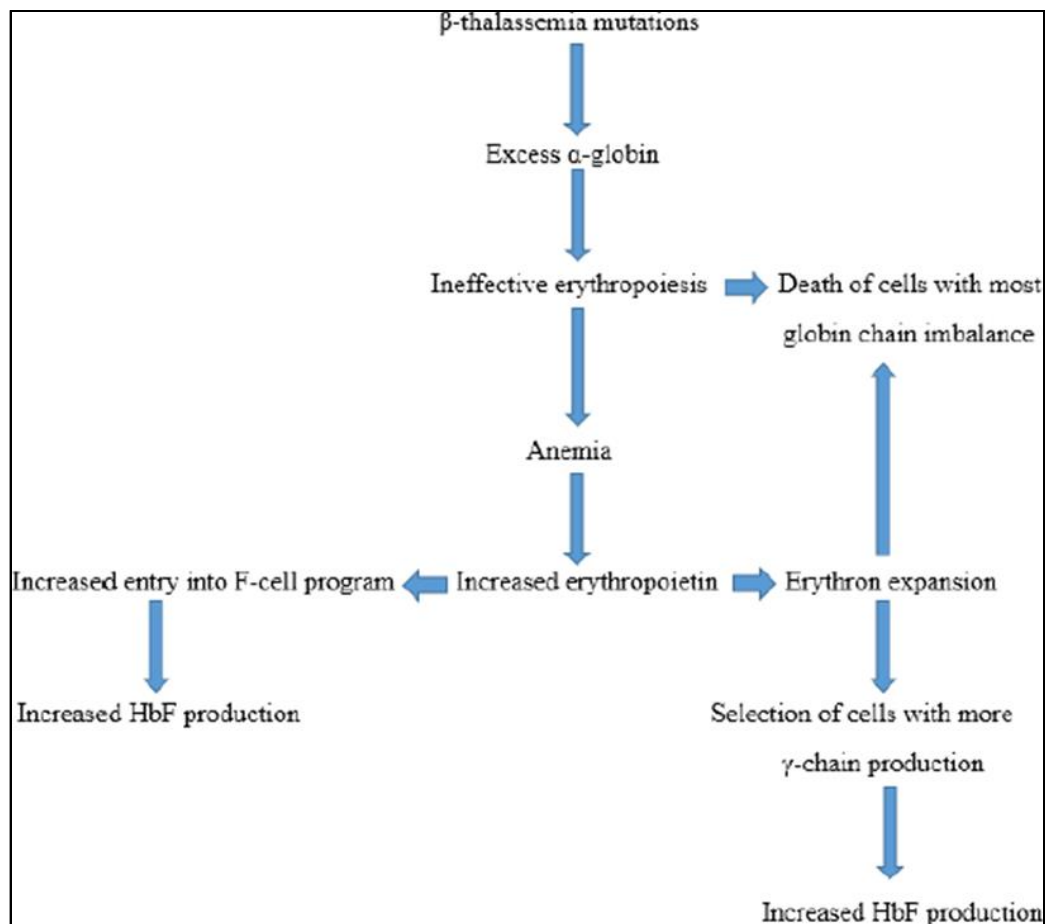


Figure 6. Proposed mechanism for increased HbF production in β-thalassemia syndromes. (Rees et al 1999)

Therapies for HbF induction

The main role of *BCL11A* in erythroid cells is the repression of HbF production. *BCL11A* also plays an important role in non-erythroid cells including self-renewal of hematopoietic stem cells (HSCs), B-cell maturation, development of the central nervous system and possibly other tissue including breast cells, dendritic cells, and pancreatic cells (Vinjamur et al 2017).

Gene therapy approaches based on hemoglobin-switching consist of modifying a patient's hematopoietic stem cells ex-vivo followed by autologous transplantation, and include:

- Genetic disruption to selectively decrease *BCL11A* expression in erythroid cells through genome editing of the intronic *BCL11A* erythroid enhancer
- Targeting *BCL11A* mRNA for degradation; using an erythroid selective shRNA lentiviral transduction
- Genetic disruption through genome editing the HBG proximal promoters to mimic HPFH mutations
- Forced looping of the LCR to the HBG promoters by lentiviral transduction of an LDB1-zinc finger fusion protein

The best available HbF-inducing medication is hydroxycarbamide, but its mechanism of action is not well understood. It therefore needs careful monitoring as patients show varying response. Advances in molecular understanding of hemoglobin switching coupled with new chemical biology approaches might enable direct targeting of HbF repressive factors like *BCL11A* and *ZBTB7A* (Vinjamur et al 2017).

Aims & Objectives

1.12.1 Aims

This study aims to evaluate whether genetic variability at a *BCL11A* locus influences HbF levels in HbE/ β -thalassemia patients in Bangladesh.

1.12.2 Specific Objectives

- To determine the frequency of single nucleotide polymorphisms (SNPs) at the *BCL11A* repressor for HbF production among HbE/ β -thalassemia patients.
- To determine the association of *BCL11A* SNP rs766432 with increased production of fetal hemoglobin among patients with

HbE/ β -thalassemia.

- To determine the effect of rs766432 on the hematological parameters in HbE/beta-thalassemia patients.
- To determine the effect of rs766432 on age at first transfusion and transfusion intervals among patients with HbE/ β -thalassemia.

CHAPTER-02

MATERIALS AND METHODS

Study Design

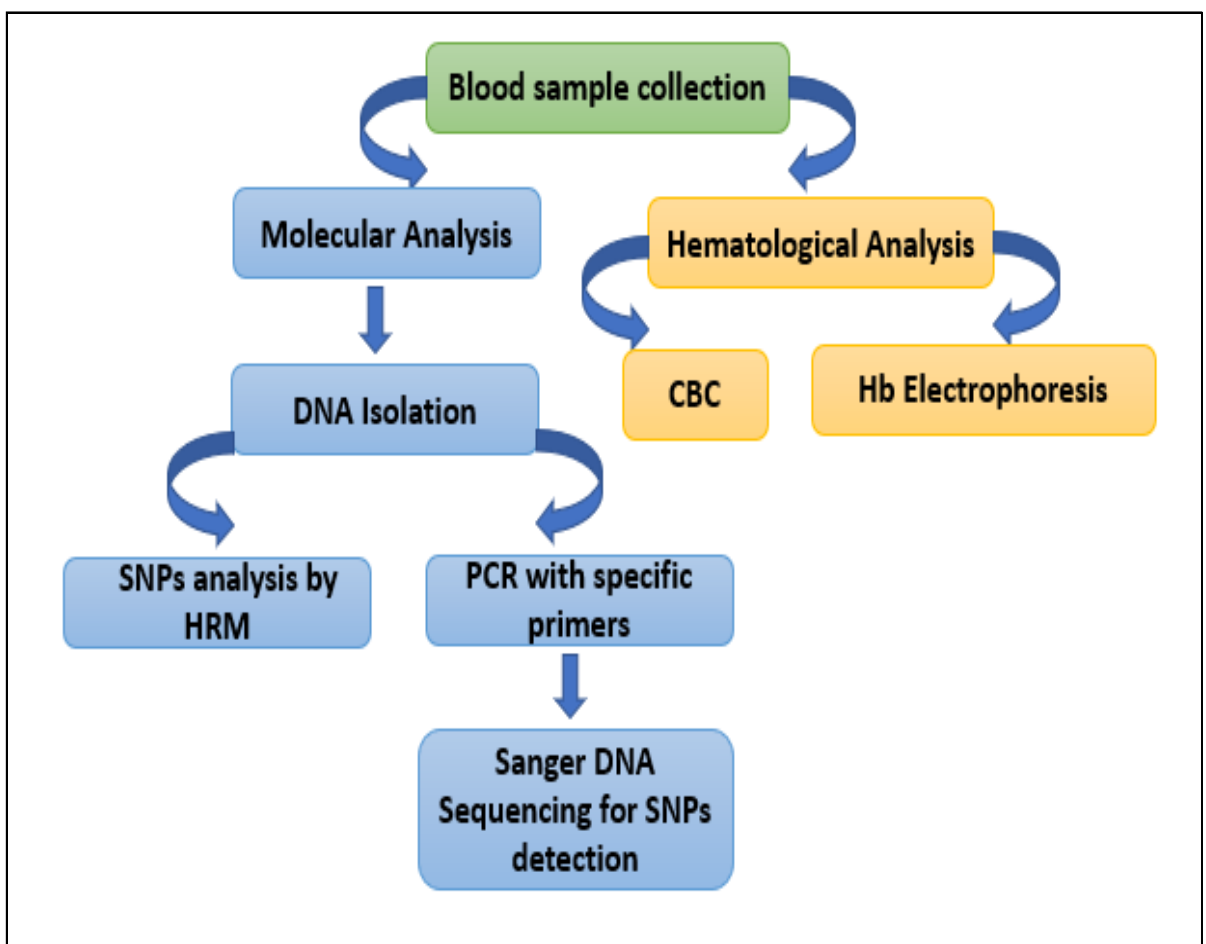


Figure 7. Workflow of the study

Study site and ethical approval

All samples were collected with the consent of patients from Thalassemia Samity Hospital and their legal guardians. All hematological, molecular, and genetic laboratory experiments were conducted in the Biosafety Level 2 (BSL-2) facilities of the Genetics and Genomics Laboratory at the Institute for Developing Science and

Health Initiatives (ideSHi). Ethical approval was received from the Bangladesh Medical Research Council (BMRC).

Sample collection site

Samples were collected with full consent of the hospital authority from September 2018 to April 2019. Patients filled up consent forms before participating in the study, with legal guardians providing written consent in the case of child patients.

Study participants

A total of 63 HbE/ β -thalassemia patients participated in the study, and blood samples were also collected from 50 unrelated healthy individuals for comparison. All participants were clinically assessed by physicians/doctors. Data on age at first blood transfusion, splenectomy condition, transfusion interval, steady state hemoglobin status, and so on was also obtained using a structured questionnaire.

Specimen collection and storage

Approximately 5 ml of peripheral blood was collected from study participants prior to blood transfusion. Portions of the blood samples (~2ml) were placed in a mix of EDTA (Ethylene Diamine Tetraacetic Acid) and BD VacutainerR (Becton Dickinson, Franklin Lakes, NJ, USA) for Hb electrophoresis analysis. The remaining blood was placed in a separate EDTA-BD VacutainerR mix for genetic analysis. These two separate EDTA-BD VacutainerR mixes were then stored in a cooler box (2-8°C) and immediately transported to ideSHi (Institute for Developing Science and Health Initiatives) for laboratory analysis. On arrival, the samples to be used for genetic analysis were stored in a freezer (-70°C) whereas the samples to be used for complete blood count (CBC) and hemoglobin electrophoresis analysis were stored in a 2-8°C refrigerator.

Hematological Analysis

Complete Blood Count (CBC) analysis

For complete blood count analysis, EDTA containing blood was analyzed using a Sysmex kx-21 (Sysmex Corporation, Kobe, Japan) following manufacturer

instructions. Hematological parameters including hemoglobin, WBC, RBC, PLT, MCV, MPV, HCT, MCH, MCHC, PCT%, Lym%, Mon%, Gran%, etc. were obtained from the CBC analysis and out of these, only data on MCV, MCH, RDW and hemoglobin levels were used in this study.

Hemoglobin Electrophoresis

Hemoglobin electrophoresis is an established technique used routinely in clinical laboratories to screen samples for hemoglobinopathies (hemoglobin variants, SCD and thalassemia). Hemoglobin electrophoresis was performed with the use of Capillarys 2 Flex-Piercing (Sebia, France). Analysis was performed using a Capillarys Hemoglobin (E) kit.

Relative quantification of individual hemoglobin fractions such as HbA, HbF and HbA2 was performed automatically following the analysis. High-resolution separation of the major hemoglobin variants (HbS, HbC, HbD and HbE) was performed using these instruments, and it is expected that HbA2 and HbF were accurately quantified by use of the hemoglobin kit.

Molecular Analysis

DNA Isolation

Genomic DNA was extracted from whole blood using a QIAGEN FlexiGeneR DNA kit (QIAGEN, Hilden, Germany).

Principle

Blood cells are lysed by adding lysis buffer to pellet cell nuclei and then making use of centrifugation. A denaturation buffer containing QIAGEN Protease and chaotropic salt is then used to efficiently remove contaminants such as protein. DNA is precipitated with 100% isopropanol, washed with 70% ethanol, and the resulting pellet is dried at room temperature and suspended in nuclease-free water.

Materials & Equipment

- Micro-centrifuge
- Water bath

- Vortex machine
- Biosafety cabinet (Level II)
- 1 ml Pipettes
- 100-200 μ L Pipettes

Reagents

- 100% isopropanol
- 70% ethanol
- Buffer FG1 (FlexiGene DNA Kit)
- Buffer FG2 (FlexiGene DNA Kit)
- Nuclease-free water (instead of Buffer FG3)
- QIAGEN Protease (FlexiGene DNA Kit)

DNA isolation procedure

- 500 μ L of FG1 buffer was placed in a 1.5 mL micro-centrifuge tube and 200 μ L of whole blood was mixed with the buffer by inverting the tube 5 times.
- The tube was centrifuged for 20 seconds at 10,000 x g using a HeraeusTM FrescoTM 21 Microcentrifuge (Thermo ScientificTM, USA).
- The supernatant was discarded, and the tube was placed upside down on a clean tissue for 2 minutes, while ensuring the pellet remained in the tube.
- A denaturation buffer was prepared by mixing QIAGEN Protease with FG2 buffer in a 1:100 ratio, and care was taken to use this buffer within an hour of preparation.
- 100 μ l FG2/QIAGEN Protease buffer was added and immediately

vortexed to completely homogenize the pellet.

- After centrifugation for 3-5s, the tube was placed in a water bath at 65 C for 5 min.
- Following incubation, 100 µl isopropanol (100%) was added and then mixed using a minimum of 20 inversions until jelly-like or thread-like DNA became visible.
- The tube was then centrifuged for 3 minutes at 10,000 x g.
- The supernatant was discarded, and the tube was placed inverted on a clean tissue for 5 minutes while ensuring the pellet remained in the tube.
- With 100 µl 70% ethanol added, the tube was vortexed for 5s and centrifuged for 3 minutes at 10,000 x g.
- The supernatant was discarded, and the tube was placed inverted on a clean sheet of tissue paper for the final air drying of the pellet, while ensuring it remains in the tube.
- The pellet was air dried for a minimum of 5 minutes until all the liquid had evaporated.
- 200 µl nuclease-free water was added to dissolve the DNA pellet and the tube was stored at 40 C overnight.
- The sample was then kept at -200 C for long-term storage.

Isolated DNA quantification

DNA concentration was measured using a Nanodrop 2000. 2µl of nuclease-free water was used as a blank and 2µl of the DNA sample was dropped on the sensor, after which its optical density (OD) was measured. The concentration (ng/µl) and purity of the DNA (260/280 nm OD) were thus obtained.

Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) was used to amplify the human *BCL11A* gene using a BioRad T100TM thermal cycler (Bio-Rad, USA) and Hot Start PCR reagents from Thermo Fisher Scientific. Two pairs of PCR primers were designed at ideSHi laboratory to amplify the 460 bp segments which cover the mutational hotspot of the *BCL11A* gene.

Principle

The basic principle of PCR is to amplify a specific segment of DNA using forward and reverse primers, Taq polymerase, PCR buffer, dNTPs, MgCl₂, and nuclease free water (Figure-8). In this reaction, Taq polymerase directs the synthesis of DNA from deoxynucleotide substrates on a single-stranded DNA template. Firstly, a primer is annealed to a longer template DNA, then the DNA polymerase adds nucleotides to the 5' end and elongates its 3' end to generate an extended region of double stranded DNA.

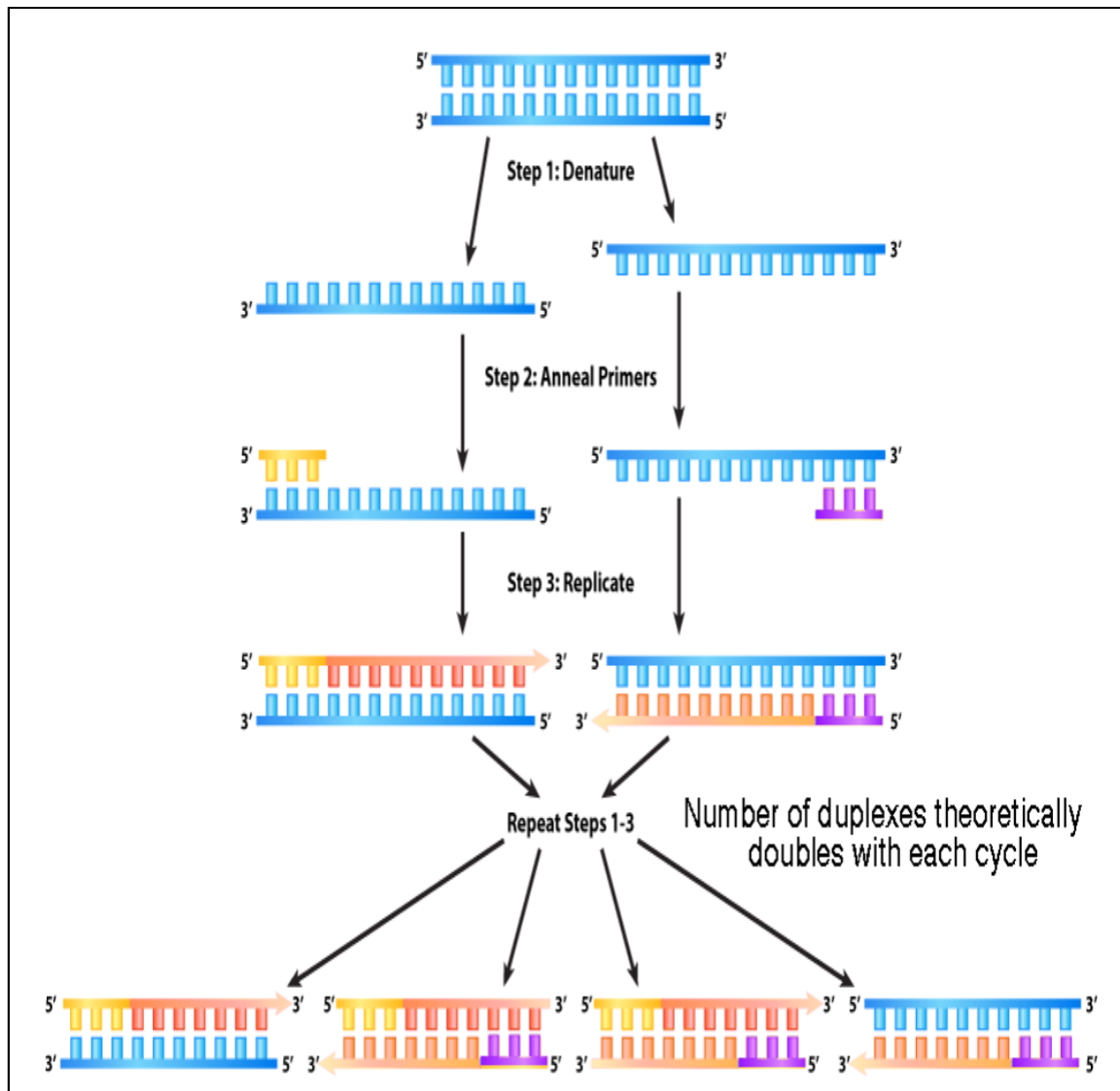


Figure 8. Polymerase Chain Reaction

Equipment & Materials

- Thermal cycler (Bio-Rad)
- Table-top spinner
- PCR cabinet
- Microcentrifuge tube
- 0.2 ml PCR tube
- 20-200 μ L Pipettes
- 1-10 μ L Pipettes

- 0.05-2.5 μ L Pipettes
- Microcentrifuge tube holder
- 0.2 ml PCR tube holder
- Tips: 0.05-2.5 μ L, 1-10 μ L, 20-200 μ L

Reagents

- Hot Start Taq DNA Polymerase
- Nuclease-free water
- 10x PCR buffer
- 25mM MgCl₂
- 10 mM dNTPs
- 10 mM Forward primer
- 10 mM Reverse primer

Procedure

Primers

The following pair of primers, designed previously by other research groups at ideSHI, was used in this study to amplify the 460 bp for the rs766432 segment of the *BCL11A* gene. Specificity and product length were checked for the amplified products of the primers. Primers, primer sequence (5'-3'), guanine-cytosine percentage (GC%), product length and primer T_m are shown in table-1 below.

Table 1. List of primers used for amplification of *BCL11A* gene

Primer (5' > 3')	Length (bp)	T _m	GC Content (%)	Product Length (bp)
<i>BCL11A</i> _rs766432_Seq_F: TGGGGGTTTCAGTGGTTAGAAGG	22	61.63	54.55	460
<i>BCL11A</i> _rs766432_Seq_R: CCAAATGCTCTGCTTATGGTGG	22	59.90	50.00	

PCR Reaction

Human *BCL11A* gene-specific PCR amplification was done using a BioRad T100TM thermal cycler (Bio-Rad, USA) and Qiagen HotStarTaq DNA Polymerase (Qiagen, Hilden, Germany). Total reaction volume was 10 uL (Table-2).

The thermal cycler was programmed to maintain the following thermal cycling conditions:

- Initial denaturation at 94 °C for 15 minutes
- 34 cycles of denaturation at 94⁰ °C for 15 seconds, annealing at 62⁰ °C for 30 seconds and extension at 72⁰ °C for 20 seconds
- final extension at 72⁰ °C for 5 minutes

Following PCR reaction, the instrument-maintained 12⁰C until shutdown (Figure-9).

Table 2. Composition of PCR reagent for amplification of *BCL11A* gene using rs766432 primers

Components	Amount (μL)
10X PCR Buffer	1.0
10 mM dNTPs	0.4
25 mM MgCl ₂	0.3
10mM Forward primer	0.3
10mM Reverse primer	0.3

Components	Amount (μL)
Hot Start Taq polymerase	0.05
Nuclease-free water	Up to 9
Master mix	9.0
Template DNA	1.0
Total reaction volume	10.0

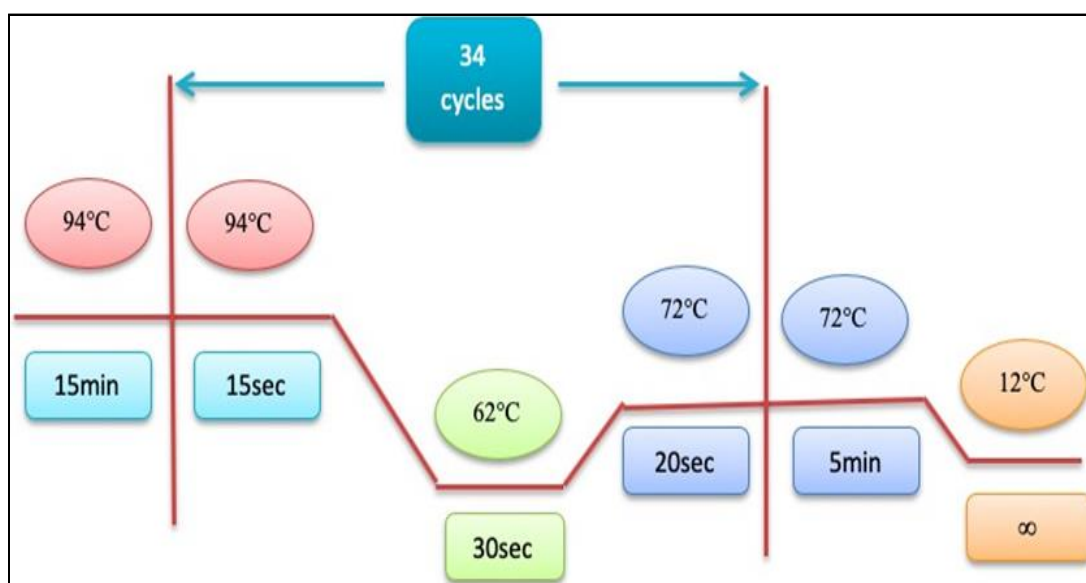


Figure 9. Thermal cycling conditions for PCR of BCL11A (rs76) polymorphic region

Agarose Gel Preparation for Gel Electrophoresis

Principle

DNA is a negatively charged molecule as a result of its phosphate groups, and at a neutral or alkaline pH, PCR products give rise to a uniform negative charge. Here, a buffer is used to maintain a constant negative charge. Tris-Acetate EDTA (TAE) can be used to obtain sharp bands and additionally prevent microbial growth over long periods of time. An electrical field is applied during electrophoresis, and in this field, the negatively charged DNA fragments move towards the positively charged electrode (anode) through the pores of the gel. The molecular weight of a DNA molecule is inversely proportional to the velocity of movement, meaning that smaller molecules travel faster than larger molecules and allow for visual differentiation.

Equipment and reagents

Equipment

- Electric balance
- Micropipettes
- Gel Electrophoresis Apparatus (Bio-Rad, USA)
- Gel Doc™ XR + (BioRad, USA)

Reagents

- Agarose powder
- 1X TAE buffer
- Gel Red (Intercalating nucleic acid fluorescent dye)
- Bromophenol Blue (Loading dye)
- DNA ladder (Marker DNA solution)

Preparation of Gel and Gel Run of PCR Products

- Preparation of 1.5% agarose gel: 1 g of agarose (ultrapure, Invitrogen, USA) was dissolved in 100 ml 1X TAE buffer (Tris-Acetate EDTA) by heating in an oven for 30 seconds.
- After cooling down slightly at room temperature, 2µl of Gel Red was added to the gel mixture.
- Once the liquid gel was lukewarm, this gel mixture was poured into a casting tray (taking care to avoid bubble formation) and a comb was set carefully in the gel. After that, the gel was allowed to solidify at room temperature.
- The comb was then removed to leave behind wells, and the gel was placed in an electrophoresis chamber filled with 1X TAE buffer.

- 2µl of loading dye was mixed with 3µl of PCR products and gently loaded into the wells.
- The final well was filled with a 1 kb+ DNA ladder as reference for the band size of PCR amplicons, and electrophoresis was run at 100 volts for 30-45 minutes.
- The gel was observed on the Gel Doc™ system (Bio-Rad, USA) under UV light and separate DNA bands were recognized (Figure-10).

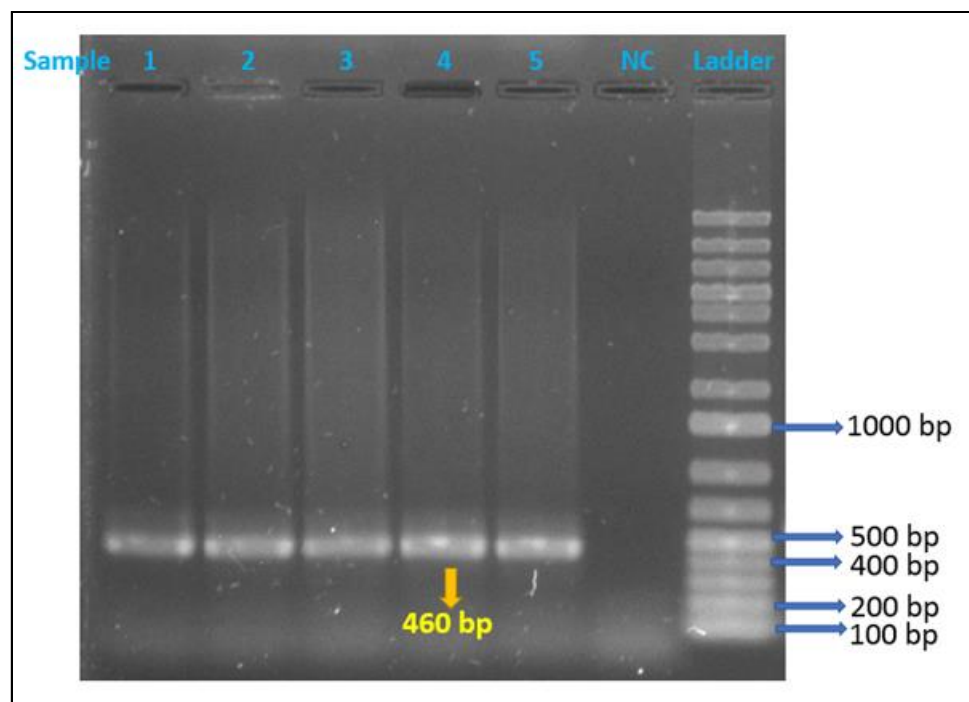


Figure 10. PCR products of BCL11A rs766432 are visualized on 1.5% agarose gel

PCR Product Purification

Principle

PCR purification involves isolating DNA from other mixed reagents. Firstly, a binding buffer absorbs DNA into the spin column membrane in an acidic solution. This buffer removes all DNA binding proteins and contaminants from PCR products as the chaotropic salts in it can destabilize hydrophobic interactions and disrupt protein structures. High salt concentration increases binding buffer affinity and stronger interaction with the column membrane simultaneously weakens the negative charge of DNA. Ethanol then removes dNTPs, primer dimers, salt, and

buffer while allowing DNA in the column to remain intact. Residual ethanol is removed through extra centrifugation. DNA elution is finally completed by adding nuclease-free water, as the neutral pH of nuclease-free water gives it a high binding affinity with DNA.

Equipment and reagents

Equipment

- Biosafety cabinet
- Micro-centrifuge tubes
- Water bath
- NanoDrop DNA quantitation system
- Refrigerator (-20 °C)

Materials

- Pipettes
- Tips
- Microcentrifuge tubes

Reagents

- Mini Elute PCR purification Kit
- 3M Sodium acetate
- Nuclease-free water

Purification procedure

- All PCR products were spun and transferred to Eppendorf tubes.
- 5 volumes of binding buffer (PB) were added to each PCR product in each microcentrifuge tube.

- The mixture was transferred to the spin column, placed on a collection tube, and binding was allowed to happen over 1-2 minutes, after which it was centrifuged for 1 min at 13,000 rpm.
- The flow-through was discarded from the MinElute column and was again placed into the same collection tube.
- 750 μ L wash buffer (PE) was added to wash each column and was again centrifuged for 1 min at 13,000 rpm.
- The flow-through was discarded and the column was placed back in the collection tube for an additional minute of centrifugation at 13,000 rpm, completely removing all residual ethanol.
- Column contents were transferred to 1.5 ml microcentrifuge tubes.
- For elution of the PCR product, 15 μ L of nuclease-free water was applied to the center of each MinElute column membrane. These were kept at room temperature for 1 minute, centrifuged at 14,000 rpm for 2 minutes to elute the DNA, and the column then discarded.
- The concentration and purity of purified DNA was measured using a Nanodrop.
- Purified DNA was then stored at -20°C for further genetic analysis.

Sanger Sequencing

Principle

DNA sequencing requires template DNA, forward/reverse primers, DNA polymerase, dNTPs, a sequencing buffer, and ddNTPs (2',3'-dideoxynucleotide triphosphates). The reaction mixtures are subjected to a thermal cycling protocol of denaturation, annealing, and extension once these components are mixed. The dNTPs bind with complementary bases as polymerase initiates amplification. The chain extension continues with dNTPs being added constantly, but chain extension is terminated after the addition of ddNTPs as their absence of a 3'-hydroxyl group acts as a chain-terminating inhibitor. These ddNTPs are tagged with fluorescent labels of

various lengths and then separated using capillary gel electrophoresis. The fragments travel from the negative electrode to the positive electrode under high voltage due to the negative charge of as DNA, and smaller fragments move through the pores more easily than larger fragments. The fragments separate based on their molecular weight and size before reaching the electrode. An optical detection device is used to detect the characteristic wavelength emitted by fluorescently labeled ddNTPs. The emitted fluorescence is converted to a digital signal and analyzed with data collection software (Figure-11).

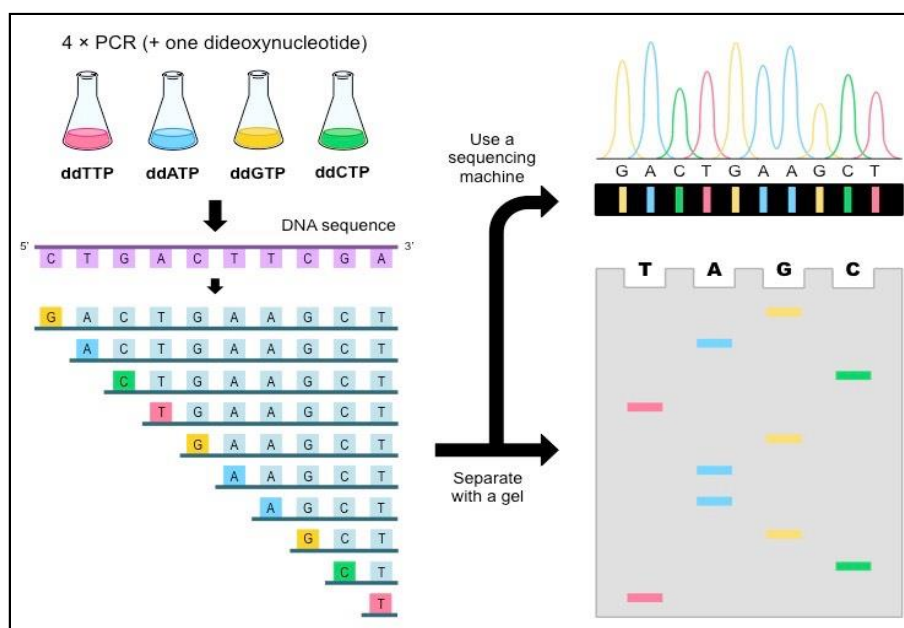


Figure 11. Sanger Nucleotide Sequencing

Procedure

Purified PCR products were sent to the Center for Medical Biotechnology (CMBT) located at the Institute of Public Health, Dhaka for sequencing. An ABI PRISM 310 automated sequencer (Applied Biosystem, USA) was used for sequencing of the purified PCR products.

Sequencing analysis

- To obtain reliable, conclusive data for mutational analysis and eliminate noisy sequences, the software Chromas Lite 2.44 was used.
- To identify mutations, non-eliminated sequences were analyzed using the Basic Local Alignment Search Tool (BLAST) from NCBI.

- The FASTA format of the sequence was aligned to detect any mismatched bases with wild type sequences which were stored in the NCBI database.
- Any mismatches found in the BLAST results were then checked in Chromas Lite to check for misreading or significant noise.

Real-Time PCR-High Resolution Melt (HRM) Curve Analysis

Principle

Real-time PCR-HRM is a powerful technique in molecular biology to detect different types of SNPs (single nucleotide polymorphisms), mutations, polymorphisms, and epigenetic differences in DNA. Firstly, the DNA is amplified using Real-Time PCR. The steps are as follows for real-time PCR - Initial denaturation, denaturation, annealing, and elongation. After the completion of real-time PCR, the amplicon DNA is melted by heating it from $\sim 50^{\circ}\text{C}$ to $\sim 95^{\circ}\text{C}$ to get different graphs for different mutated strands. The melting pattern is different for different mutations, and the resulting separate curves are seen in real time using a fluorescent dye.

Equipment & Reagents

Equipment

- Bio-Rad CFX-96 Real Time PCR machine
- Centrifuge machine
- 96 well reaction plate
- Optical plate seller
- Pipettes
- Tips
- Vortex machine

Reagents

- Type-it Master Mix (includes EvaGreen dye)
- Forward primer
- Reverse primer
- Nuclease-free water

Procedure

The real-time PCR reactions were carried out in a 96-well plate with the wells having a volume of 10 μ l. The reagents were mixed with the master-mix in a microcentrifuge tube and added in equal volumes (9 μ l) to each well, and the template DNA was then added to make a final volume of 10 μ l. The Real-time PCR Machine (CFX96 Real-Time PCR machine) was set to the following conditions for thermal cycling (Table-3):

Table 3. Thermal cycle temperatures for HRM

Steps	Temperature (°C)	Time	No. of cycle
Initial denaturation	95	2 min	1
Denaturation	95	15 sec	39
Annealing	60	30 sec	
Extension	72	20 sec	
Denaturation	95	1 min	1
Renaturation	60	1 min	1
Melting	60-95	Increment of 0.1°C per 10 sec	1

Once the real-time PCR-HRM was complete, the resulting data was analyzed with the software Precision Melt Analysis TM (BioRad). To get accurate results for cluster detection, the curve shape sensitivity was set to 100% and the T_m difference threshold was set to 0.1. Normalized and temperature-shifted views were used to successfully detect SNPs in HRM analysis.

Figure 12 is a chromatogram obtained from this analysis. It displays the substitution mutation where C is substituted by A, leading to the homozygous AA genotype. CA and AA genotypes were identified in this manner, and data were collected on the frequency of both genotypes in the DNA samples. The resulting data were used for statistical analysis.

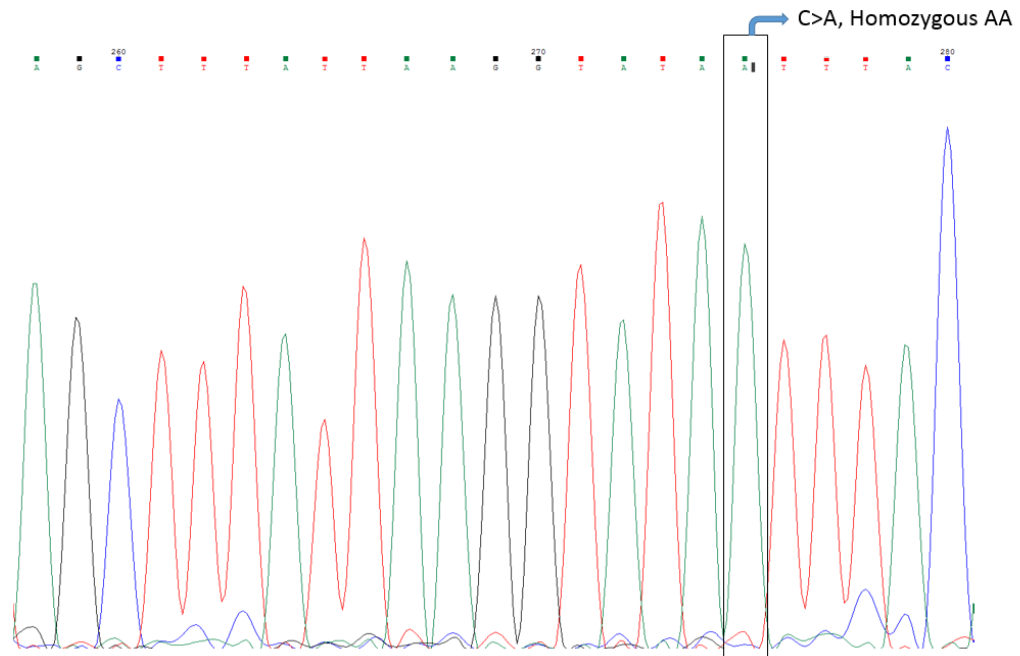


Figure 12. Sequence chromatogram showing the SNP C>A results in Homozygous major AA in BCL11A rs766432

Statistical Analysis

GraphPad Prism version 7 (GraphPad Software, La Jolla California USA) was used to perform a two-tailed T-test, one way ANOVA test, and correlation analysis on the data obtained. The data for the different variables was entered with the appropriate units, such as number (N), percentage, and mean $\pm$ standard deviation (SD). For the purpose of this analysis, a probability value (P) below 0.05 and correlation (r) near or above 0.5 was taken to be statistically significant.

CHAPTER-03

RESULTS

Demographic Information

Sixty-three HbE- β -thalassemia patients were enrolled in this study based on presence or absence of hematological parameters, hemoglobin electrophoresis data, age at first transfusion and transfusion interval data. Ages ranged from 8 years to 51 years and the Mean $\pm$ SD age was 19.66 $\pm$ 10.22 years. Among the study participants, there were 37 male patients and 26 female patients comprising 58.73% and 41.27% respectively. The Mean $\pm$ SD age among male and female patients were 26.11 $\pm$ 9.01 years and 26.18 $\pm$ 9.00 years respectively (Table-4).

Table 4. Demographic status of the HbE- β thalassemia patients

	Male	Female	Both sex
No. and % of patients	37 (58.73%)	26 (41.27 %)	63 (100.00%)
Age range (years)	8 – 51	-	-
Age (Mean $\pm$ SD) (years)	26.11 $\pm$ 9.01	26.18 $\pm$ 9.00	19.66 $\pm$ 10.22

Also, a total of 50 healthy individuals, 27 (55.0%) male and 23 (45.0%) female with a Mean $\pm$ SD age of 21.64 $\pm$ 3.31 years were recruited in this study as a control group.

Hematological status of the study patients

Hematological parameters such as Hemoglobin Concentration (HGB), Mean Corpuscular Volume (MCV), Mean Corpuscular Hemoglobin (MCH) and Red Cell Distribution Width (RDW) for each blood sample were obtained from a complete blood count (CBC) report (Table-5).

Table 5. Hematological status of the HbE- β thalassemia patients

Parameters	Results
Hemoglobin (g/dL)	7.45 $\pm$ 1.33
MCV (fL)	69.11 $\pm$ 7.54
MCH (pg)	21.68 $\pm$ 3.00
RDW (%)	25.93 $\pm$ 7.91
HCT (%)	23.48 $\pm$ 4.06

Hemoglobin status of the study patients

Hemoglobin variants such as HbA, HbA2, HbF and HbE were detected in each sample using a hemoglobin electrophoresis test (Table-6).

Table 6. Hemoglobin fractions among the study patients

Hb fractions	Absolute value (g/dL) (Mean $\pm$ SD)
HbA	3.76 $\pm$ 2.83
HbA2	0.28 $\pm$ 0.11
HbF	1.22 $\pm$ 1.55
HbE	1.96 $\pm$ 1.40

PCR results targeting *BCL11A* rs766432 among the study patients

A number of samples were selected at random for use as reference samples. Conventional PCR followed by agarose gel electrophoresis was performed to amplify and detect specific PCR products, and Sanger sequencing was done to identify polymorphic alleles. For SNP rs766432 of the *BCL11A* gene, the PCR product length was 460 bp. A 1 kb+ ladder was used to determine the sizes of base pairs.

Frequency of *BCL11A* rs766432 polymorphism among the study participant

In this study, heterozygous minor allele CA and homozygous major allele AA were found in SNP rs766432 of *BCL11A*.

Allele frequency among HbE/beta thalassemia patients

For SNP rs766432 of *BCL11A*, the genotype most commonly observed in Bangladeshi thalassemia patients was the homozygous genotype AA, which was found in 75% of study participants, whereas the CA genotype was observed in only 25% of participants (Figure-13).

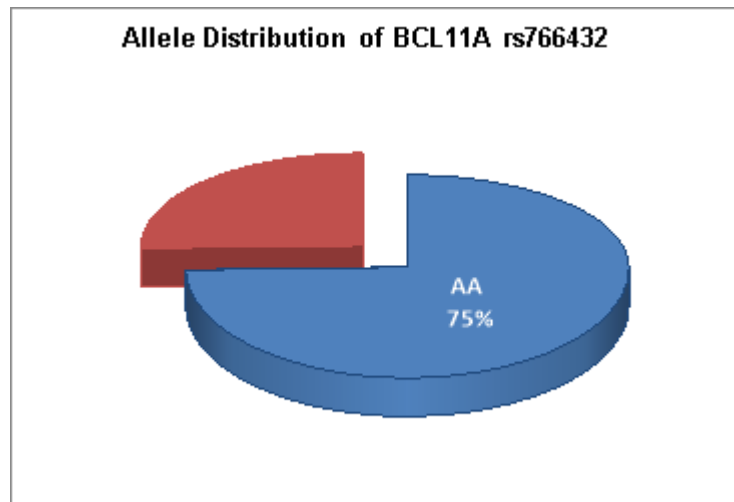


Figure 13. Allele Distributions of BCL11A rs766432 among HbE/beta thalassemia patients

Allele frequency among healthy Individuals

50 healthy individuals were recruited as the control group. All of them had completely normal hematological and hemoglobin parameters. For this control group, the genotypes found in SNP rs766432 were 74% homozygous major allele AA and 26% heterozygous minor allele CA (Figure-14).

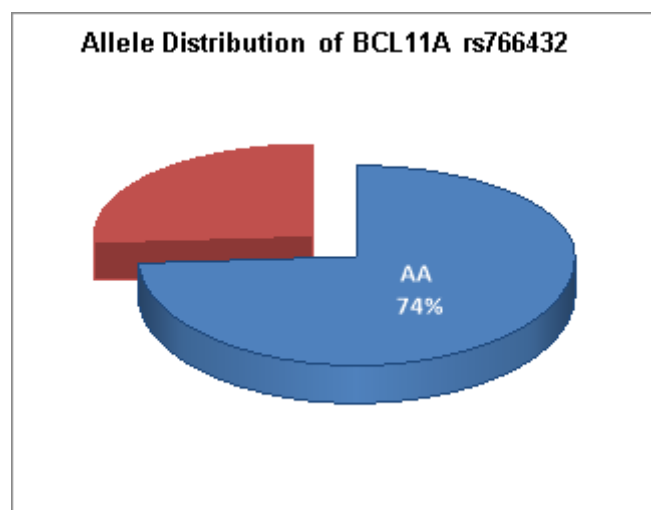


Figure 14. Allele Distributions among Healthy Individuals of BCL11A rs766432

Comparative analysis of Hb electrophoretic indices between CA and AA alleles of *BCL11A* rs766432 among the study patients

A T-test was used to compare Mean $\pm$ SD values of hemoglobin variants in study patients for alleles CA and AA in *BCL11A* rs766432. The results are shown in table-7 and figures 15-18. There were statistically insignificant differences in HbA, HbA2, HbF and HbE levels when compared between the two alleles in *BCL11A* rs766432.

Table 7. Comparison of Hb variants between two allele *BCL11A* rs766432 among the study patients

Hb variants (g/dL)	<i>BCL11A</i> rs766432		p-value
	CA	AA	
HbA	3.24 $\pm$ 2.27	3.13 $\pm$ 2.27	0.6804
HbA2	0.31 $\pm$ 0.09	0.31 $\pm$ 0.09	0.2621
HbF	1.29 $\pm$ 1.29	1.38 $\pm$ 1.35	0.1528
HbE	2.45 $\pm$ 1.20	2.50 $\pm$ 1.20	0.3234

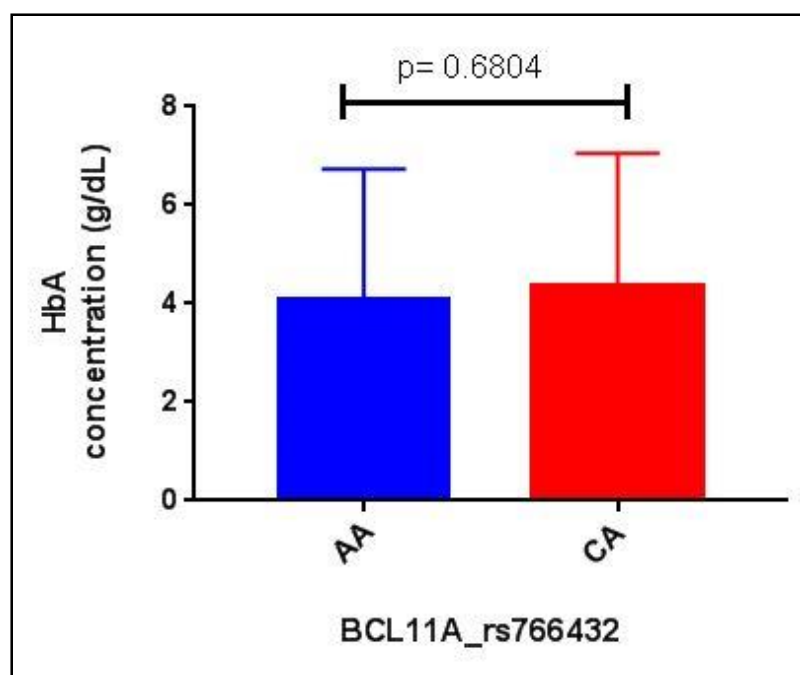


Figure 15. Comparison of HbA between two allele of rs766432

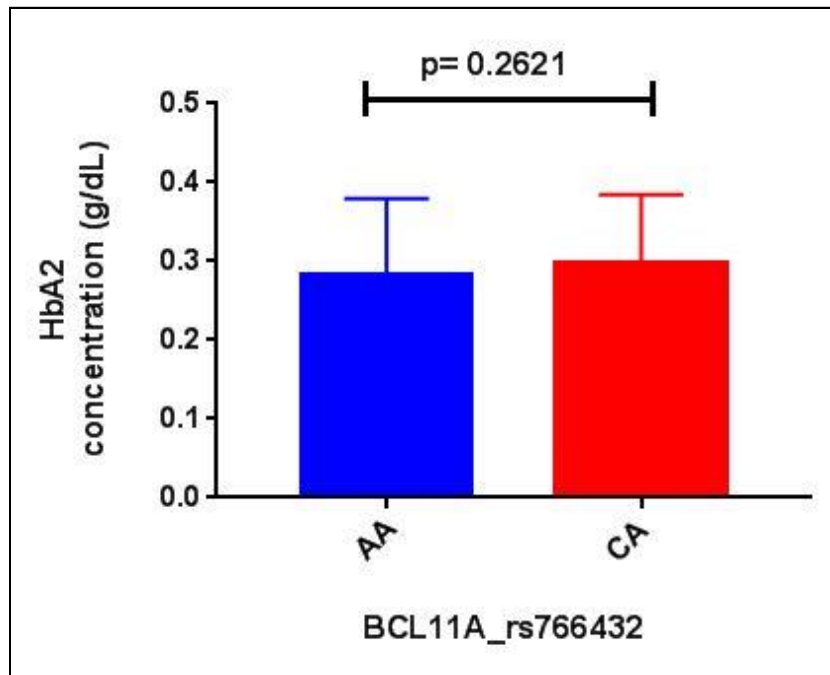


Figure 16. Comparison of HbA2 between two allele of rs766432

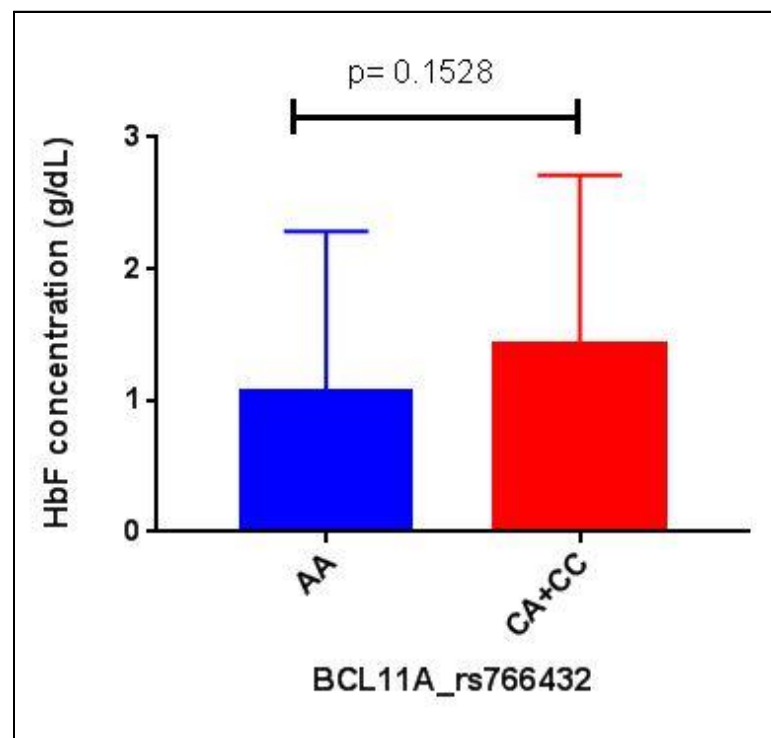


Figure 17. Comparison of HbF between two allele of rs766432

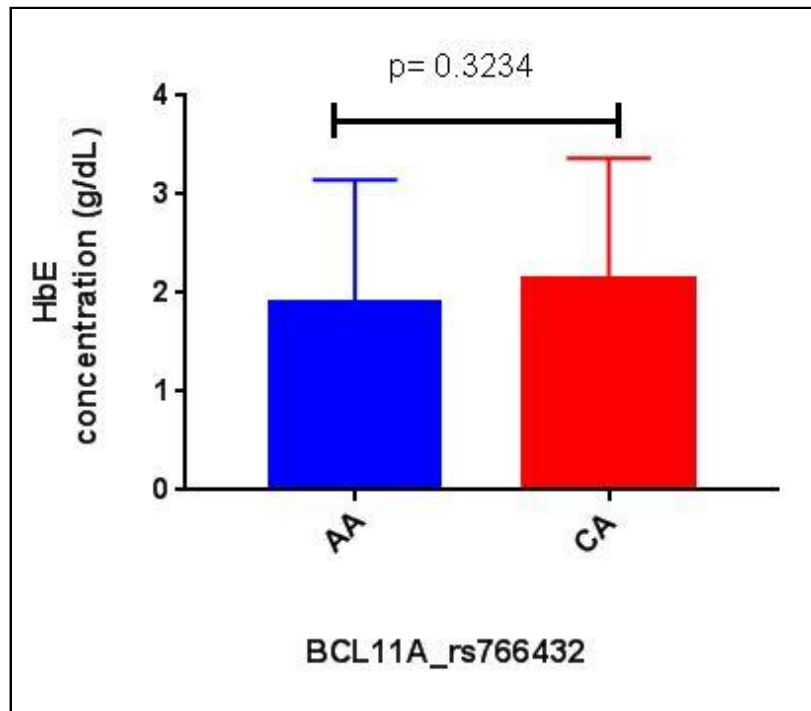


Figure 18. Comparison of HbE between two allele of rs766432

Correlation of the SNP rs766432 in *BCL11A* gene with age of first blood transfusion and transfusion intervals

Table-8 and Figure-19 depict that the average ($\pm$ SD) age of first blood transfusion for CA and AA were 98.15 ± 68.59 months and 82.65 ± 8.15 months respectively. The average ($\pm$ SD) intervals of blood transfusion were 85.29 ± 75.63 days and 63.04 ± 76.43 days for CA and AA genotype respectively.

However, neither the age at first blood transfusion nor the transfusion interval was significantly higher in patient who had rs766432 polymorphism in comparison to the patient group who had no polymorphism. Even though the mean value ($\pm$ SD) of transfusion interval demonstrates that the participants with CA genotypes had less disease severity, there is no statistically significant difference found in transfusion interval ($p=0.0582$).

Table 8. Correlation of SNP rs766432 of BCL11A gene with age of first transfusion and transfusion interval

Characteristics	BCL11A rs766432 genotype		P value
	CA (n=16)	AA (n=47)	
Age of first transfusion, month (Mean±SD)	98.15±68.59	82.65±82.15	0.226
Transfusion interval, days (Mean±SD)	85.29±75.63	63.04±76.43	0.0582

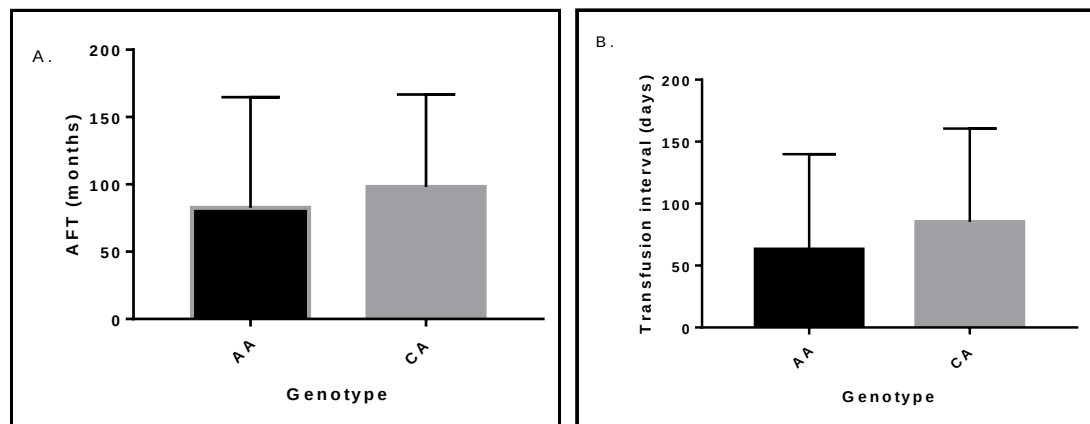


Figure 19. A Graphical Representation of the Mean±SD of (A) Age of first blood transfusion (months) (B) Blood transfusion interval (days) of different rs766432 of BCL11A genotypes

CHAPTER-04

DISCUSSION

Thalassemia, a group of genetic disorders, occurs mainly due to defective formation of the globin chain in the hemoglobin moiety of RBC. This specific type of blood disease results in excessive destruction of red blood cells which in turn leads to anemia. In this disease, RBC breakdown occurs at an early stage due to the abnormal globin chain's inability to protect the RBC from oxidative stress. The resulting destruction of RBC leads to the production of bilirubin, which is normally metabolized in the liver and disposed. In thalassemia, RBC is destroyed at such a high rate that it exceeds the liver's capacity to metabolize the excess bilirubin. Most α -thalassemia cases are caused by deletion mutations whereas β -thalassemia is caused by point mutations. The heterozygous globin gene mutation may cause anemia, and β -thalassemia major or intermedia may require blood transfusions for patient survival. Blood transfusions pose major difficulties for those afflicted, but prenatal diagnosis and genetic counseling are available to prevent thalassemia.

The prevalence of beta thalassemia is highest in South-East Asian regions. Approximately 21,693 people in South-East Asia are afflicted by β -thalassemia each year. In addition, 42,409 people are annually afflicted by β -thalassemia worldwide. Around 12% of children receiving transfusions are born with transfusion dependent β -thalassemia. HbE/ β -thalassemia is an inherited hemolytic disease highly prevalent in Bangladesh, and it has been observed that carrier rates for HbE and β -thalassemia traits are very high (8.68% and 2.24% respectively) in our population (Noor et al 2020).

According to genome-wide association studies, the three major loci associated with increasing hemoglobin F levels in HbE/ β -thalassemia patients are HBS1L-MYB, Xmn1 and *BCL11A*, which are located at chromosome 6q23.3,

promoter gene Gamma 2 and chromosome 2p16. *BCL11A* represses HbF production, which emphasizes the importance of heritable modifiers. In analyzing exome-sequencing data from healthy individuals, *BCL11A* is seen to be among the 5% of genes with the highest evolutionary constraints, so emphasis is placed on studying protective genetic variations. *BCL11A* is also expressed in hematopoietic stem and progenitor cells, and several regulators of *BCL11A* have already been identified. Determining the role of *BCL11A* regulators in an organism's development as well as in erythroid maturation will require further research (Bauer et al 2015).

In this study, heterozygous minor allele CA and homozygous major allele AA were found for SNP rs766432 in *BCL11A*. The most common genotype of rs766432 observed in the Bangladeshi study group was the homozygous genotype AA, which was found in 75% of study participants whereas the CA genotype was only observed in 25%. For healthy individuals on the other hand, 74% homozygous major allele AA and 26% heterozygous minor allele CA were observed instead. The rs766432 allele distribution in this study followed the pattern of alleles seen in almost all of the world's populations. Various populations, including Caucasians in Europe, the Caucasian population in America, and Asian populations (e.g., China, Japan, and India), indicate that the major allele of AA has a frequency of 51–88% (Rujito et al 2016).

In HbE/ β -thalassemia patients, increased levels of HbF can serve to compensate for reductions in HbA levels. The study headed by L. Rujito found that HbF increased slightly more for β^0 alleles than β alleles in HbE/ β -thalassemia patients due to cryptic mutations. Two mechanisms help in this regard, one being increased production of the gamma globin chain and the other being an increased rate of erythropoiesis, where increased rates of erythropoiesis may increase the production of F erythrocytes.

The rs766432 variety of SNPs in *BCL11A* gene is known to be associated with increasing HbF levels. *BCL11A* is a kind of regulator responsible for the transition of HbF to HbA (S Orkin et al, 2010). In a previous study, it was found that rs766432 showed a strong relationship with HbF levels for various hemoglobinopathies in Thai, Chinese, and African American populations (Chui et al 2008). In this study,

Mean $\pm$ SD values for hemoglobin variants in patients having two alleles CA and AA in rs766432 were compared and the difference found to be statistically insignificant.

However, neither age at first blood transfusion nor transfusion interval were significantly higher in patients with the rs766432 polymorphism than in the patient group without this polymorphism. Even though the mean $\pm$ SD values for transfusion interval suggest that participants with CA genotypes had a lower disease severity, no statistically significant difference was found here either ($p=0.0582$). The average $\pm$ SD age at first blood transfusion for CA and AA were 98.15 $\pm$ 68.59 months and 82.65 $\pm$ 82.15 months respectively. The average $\pm$ SD intervals for transfusion were 85.29 $\pm$ 75.63 days and 63.04 $\pm$ 76.43 days for CA and AA genotype respectively.

These results suggest multiple genetic factors in HbE/beta thalassemia patients and emphasize the importance of thorough genotyping in providing proper genetic counseling and making more effective decisions for the management and prevention of thalassemia. Increased knowledge and understanding of the genetics of HbF regulation will support the development of innovative therapeutic targets, including the development of novel drug therapies.

CHAPTER 5

CONCLUSION

Genomic sequencing revealed the roles of *BCL11A* in γ -globin regulation, making it a clear therapeutic target for inducing HbF expression. This study reveals that the SNP rs766432 has no correlation with either the HbF levels or clinical presentation of HbE-beta thalassemia patients. Despite SNP rs766432 also being present in unrelated healthy individuals, no induction of HbF expression is seen. It is concluded that HbF induction only happens in anemic conditions.

REFERENCES

- Weatherall, D.J. and Clegg, J.B. 1996. Thalassemia-a global health problem. *Nature Medicine*. 2(8):847-849.
- George, E., Khuziah, R., 1994. Malays with thalassemia in west Malaysia. *Trop Geogr Med*, 35,123-125.
- Weatherall, D.J., 2000. Introduction to the problem of hemoglobin E-beta thalassemia. *J Pediatr Hematol Onco*, 122:551.
- Algiraigri, A.H., Kassam A. Hydroxyurea for hemoglobin E/ β - thalassemia: a systematic review and meta-analysis. *International journal of hematology*. 2017 Dec 1;106(6):748-56.
- Yang, Y., Xu, Z., He, C., Zhang, B., Shi, Y., Lui, F. Structural insights into the recognition of γ -globin gene promoter by *BCL11A*. *Cell research*. 2019 Nov;29(11):960-3.
- Sedgewick, A.E., Timofeev, N., Sebastiani, P., So, J.C., Ma, E.S., Chan, L.C., Fucharoen, G., Fucharoen, S., Barbosa, C.G., Vardarajan, B.N., Farrer, L.A. *BCL11A* is a major HbF quantitative trait locus in three different populations with β -hemoglobinopathies. *Blood Cells, Molecules, and Diseases*. 2008 Nov 1;41(3):255-8.
- Fucharoen, S., Winichagoon, P. Clinical and hematologic aspects of hemoglobin E β - thalassemia. *Current opinion in hematology*. 2000 Mar 1;7(2):106-12.
- Usman, M., Moinuddin, M., Ghani, R., Usman, S., 2009. Screening of five common Beta thalassemia mutations in Pakistani population, a basis for parental diagnosis. *SQU Medical Journal*.,9(3): 305-310.
- Modell, B. and Darlison, M., 2008. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ*, 86:480–7.
- Khan, W.A. *et. al.*, 2005. Prevalence of Beta thalassemia trait and Hb E trait in Bangladeshi school children and health burden of thalassemia in our population. *DS (Child) H J*, 21 (1):1-7.

Modell, B., 1995. Update to epidemiology of hemoglobin disorders with special references to thalassemia. Thalassemia International Federation. Internet. *Transfus Med*, 5:247-58.

Khan, W.A. Prevention is the only way to combat Thalassemia. The Daily Star Web Edition 5, 1, Available from: <http://www.rightdiagnosis.com/t/thalassemia/stats.htm> [Last accessed on 2006 Feb 27].

Shekhar, H.U., Kabir, Y., Hossain, M.M., Mesbah-Uddin, M., Khatun-E-Jannat, K., Hossain, M.S., *et. al.*, 2007. Blood transfusion mediated viral infections in thalassemic children in Bangladesh. *J Med Sci*, 7:131-5.

Orkin, S.H., Kazazian, H.H. Jr., Antonarakis, S.E., Ostrer, H., Goff, S.C., Sexton, J.P., 1982. Abnormal RNA processing due to the exon mutation of beta E-globin gene. *Nature*, 300: 768-9.

Ong-Ajyooth, S., Suthipark, K., Shumnumsirivath, D., Likidlilid, A., Fucharoen, S., Pootrakul, P., 1987. "Oxidative stress and antioxidants in beta-thalassaemia/haemoglobin E". *Journal of the Medical Association of Thailand*, 70(5):270–274.

Krishnamurti, L. Few Reports of Hemoglobin E/ β -Thalassemia in Northeast India: Underdiagnosis or Complete Exclusion of β -Thalassemia by Hemoglobin E. *Journal of pediatric hematology/oncology*. 2000 Nov1;22(6):558-63.

Olivieri, NF., Muraca, GM., O'Donnell, A., Premawardhena, A, Fisher, C., Weatherall, D.J. Studies in hemoglobin E beta thalassaemia. *British journal of hematology*. 2008 May;141(3):388-97.

Steinberg, M.H., 2011. Structure and function of normal human hemoglobins. *UpToDate*, 19.3.

Hofmann, O., Mould, R., Brittain, T., 1995. Allosteric modulation of oxygen binding to the three human embryonic hemoglobins. *Biochem J*, 306 (Pt 2):367.

Sankaran, V.G., Menne, T.F., Xu, J., *et. al.*, 2008. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor *BCL11A*. *Science*, 322:1839.

Sankaran, V.G., Xu, J., Orkin, S.H., 2010. Advances in the understanding of hemoglobin switching. *Br J Haematol*, 149:181.

Borg, J., Papadopoulos, P., Georgitsi, M. *et. al.*, 2010. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat Genet*, 42:801.

- Weathall, D.J. and Clegg. J.B., 2001. Inherited Haemoglobin Disorders: An increasing global problem. *Bull World Health Organ*,79:704-12.
- Flint. J., Harding, R.M., Boyce, A.J., Clegg, J.B., 1998. The population genetics of the haemoglobinopathies. *Baillieres Clin Haematol*,11: 1-51.
- Nancy, F.O., Zahra, P., Elliott, V., 2011. HbE/beta-thalassaemia: a common & clinically diverse disorder. *Indian J Med Res*, 134:522-531.
- Organization, W.H. et. al., 16-18 November 2007. 2008. Management of haemoglobin disorders: report of a joint WHO-TIF meeting, Nicosia, Cyprus.
- Shannon, K.L., Ahmed, S., Rahman, H., Prue, C.S., Khyang, J., Ram, M. et. al., 2015. Hemoglobin E and glucose-6-phosphate dehydrogenase deficiency and Plasmodium falciparum malaria in the Chittagong Hill Districts of Bangladesh. *Am J Trop Med Hyg*, 93(2):281–6.
- Mandal, P.K., Maji, S.K., Dolai, T.K., 2014. Present scenario of hemoglobinopathies in West Bengal, India: An analysis of a large population. *Int J Med Public Heal*, 4(4):496-99.
- Mohanty, D., Colah, R.B., Gorakshakar, A.C., Patel, R.Z., Master, D.C., Mahanta, J. et. al., 2013. Prevalence of beta-thalassemia and other haemoglobinopathies in six cities in India: a multicentre study. *J Community Genet*, 4(1):33-42.
- Datta, P., Basu, S., Chakravarty, S.B. et. al., 2006. Enhanced oxidative cross-linking of hemoglobin E with spectrin and loss of erythrocyte membrane asymmetry in hemoglobin E/beta-thalassemia. *Blood Cells Mol Dis*,37:77-81.
- Pootrakul, P., Sirankapracha, P., Hemsorach, S., et. al., 2000. A correlation of erythrokinetics, ineffective erythropoiesis, and erythroid precursor apoptosis in thai patients with thalassemia. *Blood*, 96:2606-2612.
- Jetsrisuparb, A., Sanchaisuriya, K., Fucharoen, G., Fucharoen, S., Wiangnon, S., Jetsrisuparb, C. et. al., 2006. Development of severe anemia during fever episodes in patients with hemoglobin E trait and hemoglobin H disease combinations. *J Pediatr Hematol Oncol*, 28:249-53.
- Winichagoon, P., Thonglairoam, V., Fucharoen, S., Wilairat, P., Fukumaki, Y., Wasi, P., 1993. Severity differences in beta-thalassaemia/haemoglobin E syndromes: implication of genetic factors. *Br J Haematol*, 83: 633-9.

Thein, S.L., Menzel, S. Discovering the genetics underlying foetal haemoglobin production in adults. *British journal of haematology*. 2009 May;145(4):455-67.

Bauer, D.E., Orkin, S.H. Hemoglobin switching's surprise: the versatile transcription factor *BCL11A* is a master repressor of fetal hemoglobin. *Current opinion in genetics & development*. 2015 Aug 1;33:62-70.

Rees, D.C., Porter, J.B., Clegg, J.B., Weatherall, D.J. Why are hemoglobin F levels increased in HbE/ β thalassemia? *Blood, The Journal of the American Society of Hematology*. 1999 Nov 1;94(9):3199-204.

Vinjamur, D.S., Bauer, D.E., Orkin, S.H. Recent progress in understanding and manipulating haemoglobin switching for the haemoglobinopathies. *British journal of haematology*. 2018 Mar;180(5):630-43.

Noor, F.A., Sultana, N., Bhuyan, G.S., Islam, M.T., Hossain, M., Sarker, S.K., Islam, K., Khan, W.A., Rahman, M., Qadri, S.K., Shekhar, H.U. Nationwide carrier detection and molecular characterization of β -thalassemia and hemoglobin E variants in Bangladeshi population. *Orphanet journal of rare diseases*. 2020 Dec;15(1):1-2.

Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., Luc, S. *BCL11A* enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature*. 2015 Nov;527(7577):192-7.

Rujito, L., Basalamah, M., Siswandari, W., Setyono, J., Wulandari, G., Mulatsih, S., Sofro, A.S., Sadewa, A.H., Sutaryo, S. Modifying effect of XmnI, *BCL11A*, and HBS1L-MYB on clinical appearances: a study on β -thalassemia and hemoglobin E/ β -thalassemia patients in Indonesia. *Hematology/oncology and stem cell therapy*. 2016 Jun 1;9(2):55-63.

Xu, J., Sankaran, V.G., Ni, M., Menne, T.F., Puram, R.V., Kim, W., Orkin, S.H. Transcriptional silencing of γ -globin by *BCL11A* involves long-range interactions and cooperation with SOX6. *Genes & development*. 2010 Apr 15;24(8):783-98.

Sedgewick AE, Timofeev N, Sebastiani P, So JC, Ma ES, Chan LC, Fucharoen G, Fucharoen S, Barbosa CG, Vardarajan BN, Farrer LA. *BCL11A* is a major HbF quantitative trait locus in three different populations with β -hemoglobinopathies. *Blood Cells, Molecules, and Diseases*. 2008 Nov 1;41(3):255-8.