



Article

Molecular Characterization of *HBB* Gene Variations in Beta-Thalassemia Patients from Khyber Pakhtunkhwa, Pakistan

Shahzad Ahmad ¹, Laiba Khan ², Muhammad Mustafa ², Yousaf Khan ¹, Syed Farooq Shah ³, Fuzail Ahmad ¹, Taimoor Khan ¹, Muhammad Asif Zeb ¹, Qaiser Zaman ⁴ and Musharraf Jelani ^{1,2,*}

¹ Department of Molecular Biology & Genetics, Institute of Basic Medical Sciences, Khyber Medical University, Peshawar 25100, Pakistan; shahzad.ahmad.bms@kmu.edu.pk (S.A.); yousaf.ibms@kmu.edu.pk (Y.K.); fuzailahmad.mbgphd@kmu.edu.pk (F.A.); taimoorkhan@kmu.edu.pk (T.K.); muhammadasif.ipms@kmu.edu.pk (M.A.Z.)

² Rare Diseases Genetics and Genomics, Centre for Omic Sciences, Islamia College, Peshawar 25120, Pakistan; laibakhan2k26@gmail.com (L.K.); mmustafa2004h@gmail.com (M.M.)

³ Fatimid Foundation, Peshawar 25100, Pakistan; fqshah888@gmail.com

⁴ Department of Zoology, Government Postgraduate College Dargai, Malakand 23060, Pakistan; zoologist.qaiser@gmail.com

* Correspondence: mjelani.ibms@kmu.edu.pk or mjelani@icp.edu.pk; Tel.: +92-33-3520-7053

Abstract

Background: Beta-thalassemia is a hereditary hematological illness in which beta-globin chain synthesis is missing or decreased, resulting in inefficient erythropoiesis, persistent hemolysis, and anemia. It is most frequently observed in populations with a high rate of consanguineous marriages; affordability becomes a limiting factor for families with low income. **Methods:** In this study, we optimized Sanger sequencing of the whole *HBB* gene covering all three coding exons, two introns, and 5' and 3' UTRs. Eleven families were enrolled from Fatimid Foundation Peshawar, Khyber Pakhtunkhwa, Pakistan, with a total of 16 transfusion-dependent patients. An index patient in each family was Sanger sequenced for the *HBB* gene ($n = 11$). The variants were classified as per ACMG 2015 guidelines. **Results:** Sequencing analysis revealed five homozygous pathogenic variations, including three frameshifts: c.27_28dupG (p.Ser10Valfs*14), c.17_18delCT (p.Pro6Argfs17), c.126_129delCTTT (p.Phe42Leufs*19), and two splice-sites (c.92+1G>A, and c.92+5G>C) in nine families. However, two families remained unresolved. The mutation, c.27_28dupG (p.Ser10Valfs*14), was observed in several pedigrees, indicating a probable founder effect or high allele frequency within the affected population. **Conclusions:** The majority of detected variations in this study were classified as disease-causing, which reside in the first two coding exons and intron-1 of the *HBB* gene, highlighting its functional importance in the Khyber Pakhtunkhwa population. We recommend performing Sanger sequencing of exon-1-intron-1-exon-2 of the *HBB* gene with the optimized primers of this study.

Keywords: *HBB*; Sanger sequencing; thalassemia; variant; CBC



Academic Editor: Duran Canatan

Received: 28 April 2026

Revised: 22 June 2026

Accepted: 30 June 2026

Published: 10 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Thalassemia can be defined as a disease in which the production of globin chains is absent or reduced, which causes ineffective erythropoiesis, chronic hemolysis and anemia. Thalassemia is considered a common monogenic disorder globally and has the highest frequency in the Mediterranean, Middle East and South Asian regions due to genetic blueprint and social-cultural factors such as a high rate of consanguinity [1]. Thalassemia

can be classified as alpha-thalassemia and beta-thalassemia on the basis of the affected globin chains. Three clinical forms of beta-thalassemia have been identified. These include beta-thalassemia minor (carrier state), intermedia and major (Cooley's anemia) depending on severity of clinical manifestations and blood transfusion requirement [1,2]. It is caused by the mutation(s) in the globin genes, which are inherited in an autosomal recessive mode of inheritance. An individual is known to be a carrier of this disease, having a single affected gene and usually remains symptomatically healthy, whereas individuals with two affected alleles (homozygous/compound heterozygous) exhibit a severe clinical phenotype. There is a strong correlation between consanguinity and the prevalence of Thalassemia. Due to the absence of effective screening programs, the prevalence of the carrier state has increased [3,4].

Thalassemia is a major health issue globally, affecting 5–7% of the world's population who carry the different variants of the globin gene and nearly 300,000–400,000 children are born worldwide with hemoglobin disorders and beta-thalassemia [5]. Asia is the most affected region by this genetic defect, including South and Southeast Asia and the Middle East, and in these areas, the carrier rates are as high as 10–20% in the majority due to consanguinity and insufficient screening programs [4,6]. Pakistan has one of the highest numbers of children who are suffering from Thalassemia major, with a carrier rate of 5–7%, resulting in 8000–10,000 births every year and more than 100,000 patients relying on blood transfusion, which puts a great economic burden on the health sector [7]. A higher proportion of carriers is present in Khyber Pakhtunkhwa as compared with other provinces due to the marriage cultural framework in this region. There is an urgent need to establish thorough screening and genetic counseling programs for this severe inherited disease [3].

The β -thalassemia phenotype appears when mutations occur in the *HBB* gene sequence. These mutations include point mutations, insertions, deletions, or problems with the splice site. The *HBB* gene is involved in the synthesis of the β -globin chain. There is reduced or no formation of the β -globin chain when variants occur in this gene. That is why β -globin chains are fewer in number compared with α -globin chains [8,9]. The *HBB* gene is located on chromosome 11p15.5 and consists of three exons and two introns. It encodes the β -globin subunit of hemoglobin A, which is involved in carrying oxygen throughout the body in adults. Severe β -thalassemia is caused by β^0 mutations that preclude complete synthesis of hemoglobin A [10,11].

More than 1400 different variants of the hemoglobin gene have been identified. Most of them have mutations in the *HBB* gene. These variants include missense, nonsense, frameshift, or splice-site mutations. Each influences the severity of the disease differently. The distribution of these mutations differs across different regions of the world, with certain variants being more prevalent in specific populations due to founder effects and genetic drift [12,13]. In Pakistan, the majority of β -thalassemia cases are caused by a small number of mutations. The most frequently observed are c.92+5G>C, c.27_28insG, c.126_129delTTCT, c.92+1G>A and the del619 [14]. These mutations are responsible for causing 85–90% of the cases in the population [14]. The Khyber Pakhtunkhwa population lives at the border of Afghanistan and most of the ethnic groups specially Pashto-speaking-tribes descent from Afghanistan. Overall, Pakistani populations differ in sub-ethnic groups when it comes to tribal setups or local marriage. The frequency of mutations also differs in these regions [3].

β -thalassemia causes hematological changes, including reduced levels of hemoglobin, small red blood cells and low mean corpuscular hemoglobin. The RBC count can be normal or high; a raised level of HbA₂ is a sign of β -thalassemia [15]. The severity of the β -thalassemia depends on the genotype; the β^0 mutation results in severe anemia and requires transfusions, whereas the β^+ mutation is usually present in a mild clinical form [16].

More than 300 pathogenic variants of the *HBB* gene have been reported globally. High frequencies of *HBB* mutations have been documented in communities with common consanguineous marriages, notably those in Pakistan and Saudi Arabia. Consanguinity leads to the increased burden of autosomal recessive illnesses, such as β -thalassemia. Population-specific mutation spectra have been observed, emphasizing the necessity of regional molecular characterization for successful carrier screening, genetic counseling, and prenatal diagnosis. In certain regions of Pakistan, especially in Khyber Pakhtunkhwa, the spectrum of common mutations is not characterized. There is a need for sequence-based studies to find prevalent mutations and the effects of this disease. Such research would help in the improvement of diagnosis, carrier screening, and genetic counseling. Overall, the genetic landscape of β -thalassemia in these populations remains completely undefined, which highlights the importance of further investigation into *HBB* gene variants responsible for the disease.

2. Materials and Methods

Sampling: The study was conducted from September 2025 to March 2026 at the Department of Molecular Biology and Genetics, Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan. A total of eleven families were selected to participate in the study, and one index patient in each family underwent *HBB* gene screening in collaboration with Fatimid Foundation Peshawar. All families were taken into consideration based on the inclusion criteria, considering their medical signs and symptoms and clinical history and their participation was completely voluntary. The focus of inclusion criteria was on patients having severe anemia, transfusion dependency, splenomegaly and hepatomegaly.

Ethical Approval and Informed Consent Form: Prior to starting this study, the patients had provided their informed written consent to participate voluntarily and allowed academic distribution of their findings without disclosing their privacy. The study was approved by the Institutional Bioethical Review Board (IRB) of Islamia College Peshawar (ref # 529/ORIC/ICP).

Clinical Manifestation and Pedigree Designing: After identifying the clinical manifestation, which included assessment of hemoglobin (Hb) level, mean corpuscular volume (MCV), mean corpuscular hemoglobin concentration (MCHC), white blood cell (WBC) count, transfusion dependency status, and the presence or absence of splenomegaly, eleven families were taken for *HBB* gene screening. Clinical information and demographic data of the affected members of the families were recorded using a specifically designed questionnaire. Each family was assigned a unique (*HBB* ID) for their identification. Pedigrees of each family were constructed using HaploPainter software version 1.0 [17]. The questionnaire which was used for the collection of data from the affected individuals. Informed consent or assent was obtained from the affected individuals or their guardians.

Sample Collection and DNA Isolation: 3 mL of peripheral blood sample was collected in an EDTA tube (simplex) from each patient. DNA isolation was carried out by using the EasyPure (EE101-11) Genomic DNA Kit (Transgen Co., Ltd., Beijing, China). The quantity and quality of isolated DNA were analyzed using a Nanodrop spectrophotometer (Thermo Scientific NanoDrop 2000c, Thermo Fisher Scientific, Waltham, MA, USA) and the integrity was assessed using agarose gel electrophoresis (BioRad, Hercules, CA, USA).

Complete Blood Picture: CBC was performed from EDTA blood using a Sysmex analyzer (XP 300, Kobe, Japan).

Sanger Sequencing: One index patient (affected one) having clinical features from each enrolled family was chosen for the *HBB* Sanger Sequencing, which was performed at ACGT Laboratory, Khyber Medical University, Peshawar. The genomic DNA se-

quence of the *HBB* gene was retrieved from the Ensembl database (<https://grch37.ensembl.org/index.html> accessed on 10 April 2026). Five pairs of overlapping primer sets (Supplementary Materials) were selected from the genomic sequence of the *HBB* gene to capture all three coding exons, two introns and the 5'UTR and 3'UTR. The product size of every primer set was ~500 bp and overall, the whole genomic sequencing of a 2.056 Kb region was sequenced in five different reactions. These primer sets were PCR optimized at 60 °C using 5 min primary denaturation at 95 °C, 40 cycles of 30 s denaturation at 95 °C, 30 s annealing at 60 °C, 30 s extension at 72 °C, final extension of 5 min at 72 °C. Sanger Sequencing of affected samples was carried out using Genetic Analyzer SeqStudio 6000 (Thermo Fisher Scientific, Waltham, MA, USA).

Pathogenicity assessment: To predict the disease-causing variant and to detect previously reported and novel mutations were checked with ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>, accessed on 29 June 2026), gnomAD v4.1.1 (<https://gnomad.broadinstitute.org/>, accessed on 29 June 2026) and in-silico prediction tools such as MutationTaster2025 (<https://www.mutationtaster.org/>, accessed on 29 June 2026), SpliceAI (<https://spliceailookup.broadinstitute.org/>, accessed on 29 June 2026), REVEL (<https://sites.google.com/site/revelgenomics/>, accessed on 29 June 2026), and CADD (<https://cadd.bihealth.org/>, accessed on 29 June 2026).

3. Results

A total of 15 cases were identified across 11 families representing the Pakhtun ethnic group of Pakistan (Figure 1). One affected individual was enrolled from each family ($n = 11$); both gender inclusive, with ages ranging from 6 to 26 years. Patients from 8 out of 11 families were born to consanguineous marriages, revealing an autosomal recessive pattern of inheritance, while the other four families had no consanguinity. The distribution of affected individuals across multiple family members within certain pedigrees (e.g., HBB-03, HBB-04, HBB-07, HBB-10, and HBB-11) suggests familial aggregation, supporting an inherited pattern of disease transmission. Overall, the cohort characteristics highlight a predominantly young population with a high frequency of consanguinity, which is a known contributing factor to the increased prevalence of inherited genetic disorders in this population.

The complete blood count (CBC) and clinical data of the affected individuals demonstrated a consistent pattern indicative of a severe thalassemia phenotype. Hemoglobin (Hb) levels were markedly reduced across most patients, typically ranging between approximately 3 and 7.4 g/dL, reflecting profound anemia. Red blood cell indices further supported this observation, with mean corpuscular volume (MCV) values generally decreased (around 71–86 fL), indicating microcytosis, and mean corpuscular hemoglobin (MCH) values also reduced (approximately 22–27 pg), consistent with hypochromia. Mean corpuscular hemoglobin concentration (MCHC) remained relatively within or slightly below the normal range in most individuals. White blood cell (WBC) counts showed mild variation without a consistent abnormal trend, while platelet (PLT) counts varied among patients, with some individuals exhibiting elevated levels, possibly reflecting compensatory hematopoietic responses (Table 1). Clinically, all individuals were transfusion-dependent and were classified as having a severe disease phenotype, highlighting the significant clinical burden of the disorder. Notably, the observed clinical and hematological similarities among affected individuals across different families correlate with the underlying genetic findings, where the same *HBB* gene variant was identified in five families, while the remaining families harbored variants localized within the same exon, further supporting a strong genotype–phenotype association in this cohort.

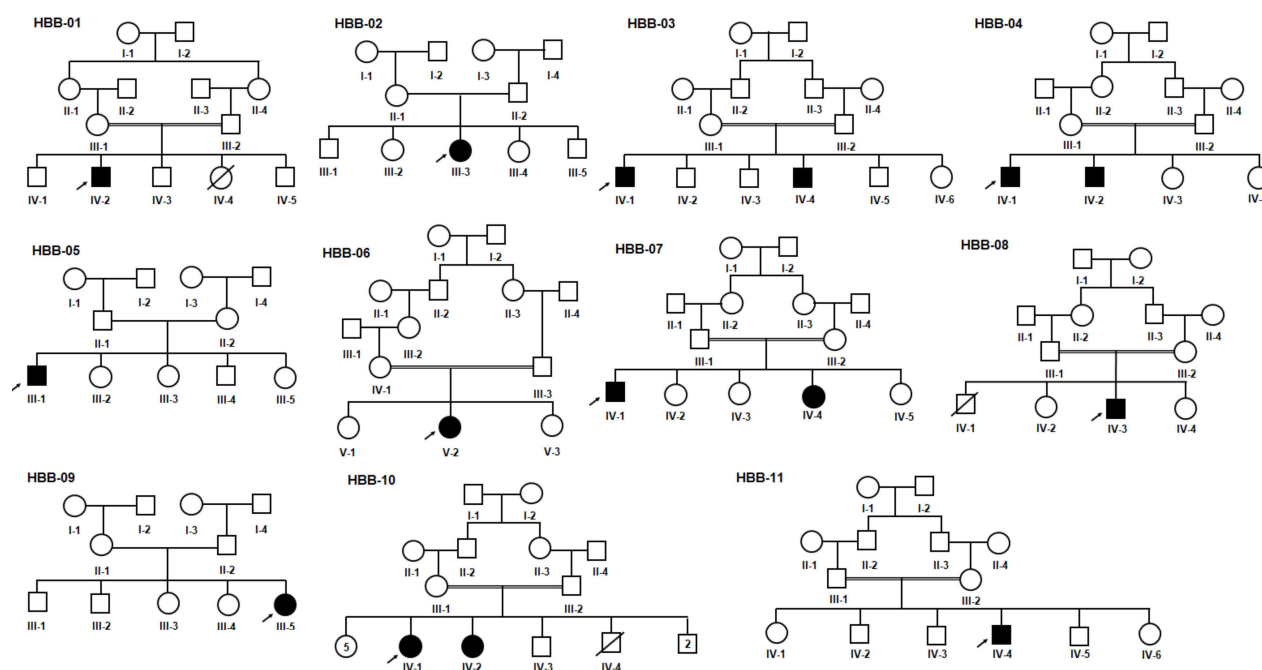


Figure 1. Pedigrees of eleven thalassemia families (HBB-01, HBB-02, HBB-03, HBB-04, HBB-05, HBB-06, HBB-07, HBB-08, HBB-09, HBB-10, HBB-11). Squares denote males, while circles denote females. The oblique signs on the symbols denotes deceased individuals. Filled squares and circles indicate affected members of the family and empty symbols indicate unaffected individuals. A double marriage line between individuals represents consanguinity; arrows indicate the index patients in each family.

Molecular analysis of the *HBB* gene identified known homozygous pathogenic variants in the nine enrolled cases (Table 2). Three pathogenic frameshifts; c.27_28insG (p.Ser10Valfs14*), c.17_18delCT (p.Pro6Argfs17), c.126_129delCTTT (p.Phe42Leufs*19), two splice-site; c.92+5G>C, c.92+1G>A, and a missense; c.9T>C variants were identified across the cohort. All the causative variants identified in these families were detected with one primer set, HBB_2F1: 5'-GCAGAGCCATCTATTGCTTA-3' and HBB_2R1: 5'-AAACATCAAGCGTCCCATAG-3' PCR Product = 540 bp. Despite repetitive target amplification of non-coding intron-2 and exon-3 and 3'UTR of *HBB*, the first PCR was not successful in families HBB-04 and HBB-07. This was indicative of a large deletion (del619) extending from intron-2-exon-3-3'UTR; however, the exact sequence breakpoint could not be identified in our analysis.

Sequencing analysis also revealed some common polymorphisms, including c.9T>C, 316-185C>T, which were declared as “benign” as per ACMG 2015 criteria [18]. The pathogenic variant c.27_28insG (p.Ser10Valfs*14) was identified in multiple pedigrees (HBB-05, HBB-02, HBB-06, HBB-03, and HBB-10), indicating a possible founder effect or high allele frequency within the population. Overall, the majority of identified variants were classified as disease-causing and localized within exon 1 or its flanking regions, reinforcing the critical functional importance of this region in the *HBB* gene (Figure 2).

Table 1. Clinical features of eleven affected individuals in thalassemia families.

Pedigree ID	Individual ID	Age at Enrollment (Years)	Gender	Consanguinity	Population Caste	WBC	HB	MCV	MCH	MCHC	PLT	Transfusion Dependent	Stages of Phenotype
HBB-01	IV-2	8	Male	Yes	Pashtun	12.5	7.5	79	27	43	712,000	Yes	Severe
HBB-02	III-3	5	Female	No	Pashtun	7.5	3	75	27	36	163,000	Yes	Severe
HBB-03	IV-1	18	Male	Yes	Pashtun	4.4	4.8	77	25	33	312,000	Yes	Severe
	IV-4	10	Male			-	-	-	-	-	-	Yes	Severe
HBB-04	IV-1	16	Male	Yes	Pashtun	5	7	77	26	34	283,000	Yes	Severe
	IV-2	14	Male			-	-	-	-	-	-	Yes	Severe
HBB-05	III-1	16	Male	No	Pashtun	57	6.4	76	22	29	1,183,000	Yes	Severe
HBB-06	V-2	9	Female	Yes	Pashtun	2.7	3.4	81	27	33	113,000	Yes	Severe
HBB-07	IV-1	20	Male	Yes	Pashtun	11.8	7.4	72	24	33	418,000	Yes	Severe
	IV-4	13	Female			-	-	-	-	-	-	Yes	Severe
HBB-08	IV-3	8	Male	Yes	Pashtun	7	5.2	76	86	34	413,000	Yes	Severe
HBB-09	III-5	11	Female	No	Pashtun	4.6	3.5	79	25	32	239,000	Yes	Severe
HBB-10	IV-1	26	Female	Yes	Pashtun	6	5.4	80	27	33	363,000	Yes	Severe
	IV-2	13	Female			-	-	-	-	-	-	Yes	Severe
HBB-11	IV-4	25	Male	Yes	Pashtun	10		8.1	71	24	33	Yes	Severe

Table 2. List of variants detected through Sanger sequencing in *HBB* gene in thalassemia patients.

Family ID	Label ID	Gene	Position	Nucleotide Change	Protein Change	Zygoty	Mutation Tester Prediction
HBB-08	08	<i>HBB</i>	Exon 1	c.17_18delCT	p.Pro6Argfs*17	Homozygous	Disease causing
HBB-05	05	<i>HBB</i>	Exon 1	c.27_28insG	p.Ser10Valfs*14	Homozygous	Disease causing
HBB-02	02	<i>HBB</i>	Exon 1	c.27_28insG	p.Ser10Valfs*14	Homozygous	Disease causing
HBB-06	06	<i>HBB</i>	Exon 1	c.27_28insG	p.Ser10Valfs*14	Homozygous	Disease causing
HBB-10	10	<i>HBB</i>	Exon 1	c.27_28insG	p.Ser10Valfs*14	Homozygous	Disease causing
HBB-01	01	<i>HBB</i>	Exon 1	c.27_28insG	p.Ser10Valfs*14	Homozygous	Disease causing
HBB-09	09	<i>HBB</i>	Intron 1	c.92+1G>A	p.?	Homozygous	Disease causing
HBB-11	11	<i>HBB</i>	Intron 1	c.92+5G>C	p.?	Homozygous	Disease causing
HBB-03	03	<i>HBB</i>	Exon 2	c.126_129delCTTT	p.Phe42Leufs*19	Homozygous	Disease causing

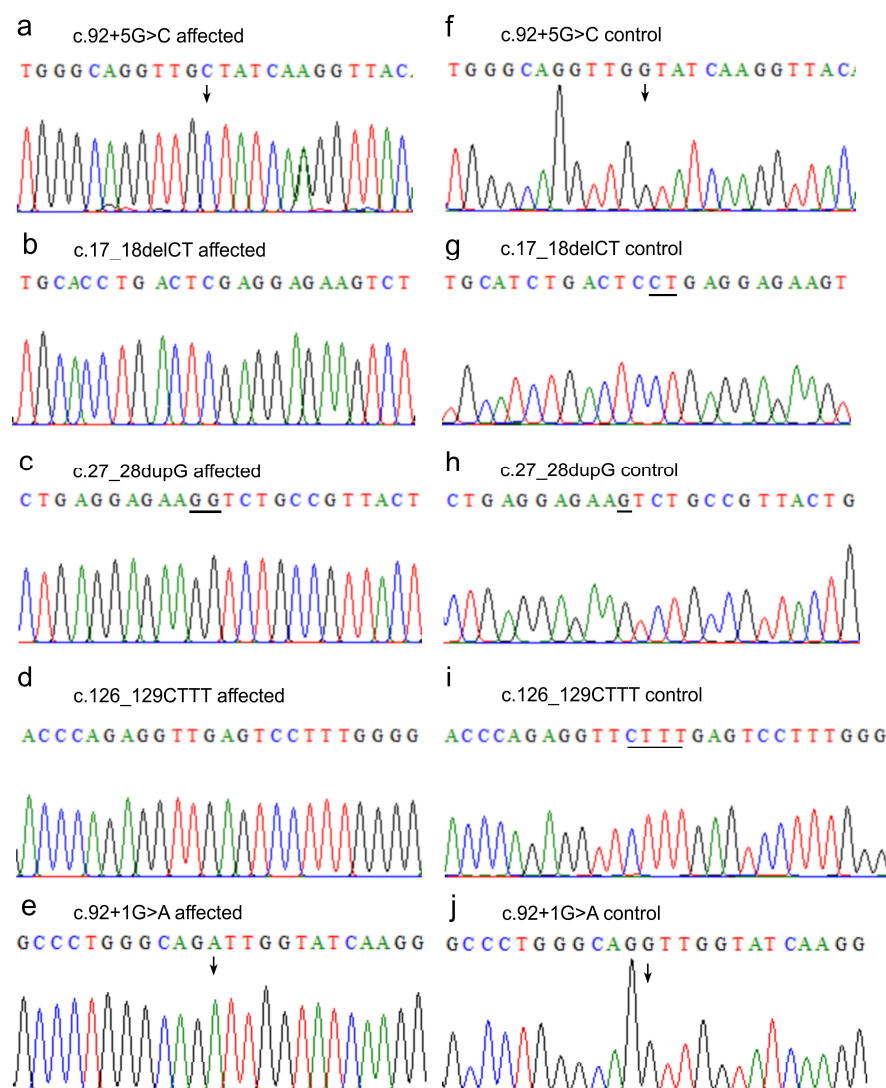


Figure 2. Sanger sequencing analysis of the five homozygous HBB variants identified in this study. (a) homozygous c.92+5G>C in family 11 (b) homozygous c.17_18delCT in family 8 (c) homozygous c.27_28dupG in families 1, 2, 5, 6 and 10 (d) homozygous c.126_129delCTTT in family 3 (e) homozygous c.92+1G>A in family 9, and (f–j) healthy control or wild type sequences. The underlined nucleotides indicate the point of deletion/insertion, while arrows indicate substitution variation in the respective sequence analysis.

4. Discussion

The high prevalence of thalassemia in Pakistan, especially in the Khyber Pakhtunkhwa province, is the result of a high rate of consanguineous marriages and multigenerational families. Most of the patients studied within the cohort were born to consanguineous marriages, while some had non-consanguineous pedigrees. Previous studies have associated *HBB*, *HBA1*, and *HBA2* genes with the thalassemia syndrome [19]. In the current study, of all eleven recruited thalassemia transfusion-dependent patients, the majority of identified variants were classified as disease-causing and localized within exon 1 or its flanking regions, reinforcing the critical functional importance of this region in the *HBB* gene. The presence of homozygous variants among the unrelated families presents a unique case of identity by descent, indicating a founder effect and connecting these families to a common ancestry. These findings are consistent with previous research on the *HBB* gene, which accounts for a significant proportion of familial thalassemia cases, as reported by Ejaz et al. [13,20].

The variant c.27_28insG p.Ser10Valfs*14 was the most frequent *HBB* gene mutation observed in our study. This mutation is reported to have the highest distribution in the Pakistani population previous study by Humaira et al. [13,21,22]. The insertion of a single nucleotide disturbs the usual reading frame, leading to a premature termination codon and a shortened, nonfunctional β -globin. Such variants are typically associated with a β^0 -thalassemia phenotype, resulting in absent β -globin chain synthesis and severe clinical appearances [23]. The preponderance of this variant in our study may reflect a regional hotspot or founder effect, as comparable results have been observed in South Asian populations [20,24]. Previous research has also revealed that frameshift mutations like these are highly associated with severe hematological indices and transfusion-dependent symptoms, particularly in homozygous or compound heterozygous individuals. These findings highlight the clinical significance of this variant and justify its inclusion in standard molecular screening and prenatal diagnostic programs [13]. Our study did not have the c.92+5G>C mutation, which is a common β -thalassemia variant in Pakistan. This could be due to regional variation or insufficient sample representation [14,25].

The c.92+1G>A variant found in family HBB-09 is a key splice-site mutation in intron 1 of the *HBB*, impacting the highly conserved donor splice site at the exon-intron boundary. This location is critical for proper RNA splicing, and substitution at the +1 site usually causes full disruption of normal splicing, resulting in abnormal mRNA transcripts or exon skipping. This mutation is usually associated with a β^0 -thalassemia phenotype, which is defined by the absence of functional β -globin chain production. Notably, this variant has been described as one of the most common mutations in earlier research from Pakistan, including the work by Ahsan H et al., further validating its regional occurrence and clinical significance [26].

Molecular analysis of the *HBB* gene identified pathogenic variants using Sanger sequencing in 9 out of 11 cases. This ~80% diagnostic yield makes Sanger sequencing ideal for the CVS, carrier, and molecular screening of thalassemia syndrome. The study presents exon 1 and its flanking splice-site regions as hotspots, which are critical for functional β -globin synthesis [10]. Specifically, the identification of splice-site mutations such as c.92+5G>C and c.92+1G>A highlights the role of aberrant mRNA processing in causing severe phenotypes [13]. These findings are supported by a recent mutational analysis of Pakistani patients, which shows that mutations in these specific genetic regions severely affect hemoglobin synthesis [11]. Our pedigrees show a substantial association between genotype and phenotype, indicating the prevalence of severe β -globin thalassemia mutations in Khyber Pakhtunkhwa, Pakistan.

The frameshift variant c.17_18delCT (p.Pro6Argfs*17) was found in the patient from family HBB-09 and is classed as pathogenic. This deletion introduces a premature stop codon, resulting in a truncated and likely nonfunctional β -globin protein, thereby contributing to the disease phenotype. Notably, this mutation has also been found recently in two more studies, further establishing its clinical significance and recurrent nature in affected groups [27–29]. A nonsynonymous variant, c.9T>C and c.315-185C>T in the deep intronic region of intron-2 were identified in our cohort and were declared as population-specific polymorphisms. The benign variants are usually not reported in molecular diagnostics; however, this variant appears in the majority of patients in the Pakhtun population. This is why it should be noted as a common polymorphism in this population.

Our study's hematological data revealed significantly lower hemoglobin (Hb) values, ranging from 3 to 7.4 g/dL, which are indicative of the TDT phenotype. These findings, along with reduced mean corpuscular volume (MCV: 71–86 fL) and mean corpuscular hemoglobin (MCH: 22–27 pg), are consistent with the global clinical markers for severe anemia [2,30]. Our findings are consistent with recent pediatric studies in the region, which

show that pediatric patients in high-prevalence areas frequently exhibit these specific microcytic hypochromic indices [5]. Furthermore, the severity of anemia found in our cohort is in close agreement with the hematological spectrum reported in previous Pakistani patient studies, emphasizing the durability of these markers as diagnostic indicators for the primary form of the disease [31,32].

The study has significant limitations. First, though the sample size in this study does not represent the whole population of beta-thalassemia patients in Khyber Pakhtunkhwa, we have seen that 9 out of 11 families were resolved for molecular diagnosis with only one primer set covering exon-1-intron1-exon-2 of the *HBB* gene. Second, the two families that were negative may be resolved with tools other than Sanger sequencing, for example, gap-PCR assay [33], to confirm their large deletion of 619 bp, including exon-3 and 3'UTR. Finally, we conclude that the restricted number of families may limit the findings' generalizability. Future studies with larger cohorts and thorough sequencing methodologies are needed to better understand the molecular spectrum of β -thalassemia.

5. Conclusions

Our findings confirm *HBB* hotspot mutations in Khyber Pakhtunkhwa, with variants such as c.27_28insG accounting for a substantial proportion of cases despite overall mutational heterogeneity. The Government of Khyber Pakhtunkhwa has already passed a law to screen families and perform carrier testing for beta-thalassemia before marriages. This genetic testing is not only efficient but also cost-effective for screening the most common variants. The five primer sets can be used stepwise to reduce cost and to improve the diagnostic yield in the families of Khyber Pakhtunkhwa, Pakistan. Hematological findings, such as anemia with microcytosis and hypochromia, corroborated the molecular results, reinforcing the genotype-phenotype association and emphasizing the clinical need for targeted *HBB* gene screening in this population. This regional clustering facilitates the creation of a cost-effective, focused *HBB*-based diagnostic panel for local screening.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/thalassrep16030014/s1>. *HBB* genomic sequence with primers sets.

Author Contributions: S.A.—wrote the manuscript; M.J.—designed experiments and reviewed the manuscript; T.K.—variant curation, reviewed results and discussion; Y.K., M.M. and L.K.—data analysis; S.A.—prepared figures, made tables; Q.Z., M.A.Z. and S.F.S.—reviewed the manuscript; F.A.—Software, formal analysis; M.J. and Q.Z.—supervised the project, Project administration, reviewed, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financially supported not supported by any funding agency.

Institutional Review Board Statement: The study was approved by the Institutional Bioethical Review Board (IRB) of Islamia College Peshawar (ref # 529/ORIC/ICP). All the safety and privacy measures were taken under the guidelines of Helsinki's Declaration 2013 standards.

Informed Consent Statement: Prior to starting this study, the patients had provided written informed consent to participate in this study voluntarily.

Data Availability Statement: The raw data files are available in Supplementary Materials.

Acknowledgments: We are grateful to the thalassemia families for their voluntary cooperation in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Taher, A.T.; Weatherall, D.J.; Cappellini, M.D. Thalassaemia. *Lancet* **2018**, *391*, 155–167. [[CrossRef](#)] [[PubMed](#)]
2. Cappellini, M.D.; Cohen, A.; Porter, J.; Taher, A.; Viprakasit, V. *Guidelines for the Management of Transfusion Dependent Thalassaemia (TDT)*; Thalassaemia International Federation: Strovolos, Cyprus, 2021.
3. Haq Eu Naeem, A.; Awan, U.A.; Khan, S.N.; Zohaib, M.; Uddin, G. High prevalence and molecular spectrum of β -thalassemia trait in Haripur and Abbottabad districts of Pakistan: Implications for regional carrier screening programs. *Egypt. J. Med. Hum. Genet.* **2025**, *26*, 167. [[CrossRef](#)]
4. Modell, B.; Darlison, M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull. World Health Organ.* **2008**, *86*, 480–487. [[CrossRef](#)] [[PubMed](#)]
5. Mohammad, S.N.N.A.I.; Iberahim, S.; Ab Rahman, W.S.W.; Hassan, M.N.; Edinur, H.A.; Azlan, M.; Zulkafli, Z. Anemia in children: A review. *Biomed. Res. Ther.* **2025**, *12*, 7837–7845. [[CrossRef](#)]
6. Colah, R.; Gorakshakar, A.; Nadkarni, A. Global burden, distribution and prevention of β -thalassemias and hemoglobin E disorders. *Expert Rev. Hematol.* **2010**, *3*, 103–117. [[CrossRef](#)] [[PubMed](#)]
7. Khaliq, S. Thalassemia in Pakistan. *Hemoglobin* **2022**, *46*, 12–14. [[CrossRef](#)] [[PubMed](#)]
8. Origa, R. β -Thalassemia. *Genet. Med.* **2017**, *19*, 609–619. [[CrossRef](#)] [[PubMed](#)]
9. Higgs, D.R. The molecular basis of α -thalassemia. *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a011718. [[CrossRef](#)] [[PubMed](#)]
10. Thein, S.L.; Rees, D. Haemoglobin and the inherited disorders of globin synthesis. In *Hoffbrand's Postgraduate Haematology*; John Wiley & Sons: Hoboken, NJ, USA, 2025; pp. 222–248.
11. Alsaleh, H.A. Differential Gene Expression Analyses in HBE/Beta Thalassaemia Patients and Their Relationship to Disease Severity. Ph.D. Thesis, Universiti Sains Malaysia, Pulau Pinang, Malaysia, 2022.
12. Hardison, R.C.; Chui, D.H.; Giardine, B.; Riemer, C.; Patrinos, G.P.; Anagnou, N.; Miller, W.; Wajcman, H. HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. *Hum. Mutat.* **2002**, *19*, 225–233. [[CrossRef](#)] [[PubMed](#)]
13. Ejaz, S.; Abdullah, I.; Usman, M.; Iqbal, M.A.; Munawar, S.; Khan, M.I.; Imtiaz, N.; Tahir, H.; Bari, M.I.; Rasool, T.; et al. Mutational analysis of hemoglobin genes and functional characterization of detected variants, through in-silico analysis, in Pakistani beta-thalassemia major patients. *Sci. Rep.* **2023**, *13*, 13236. [[CrossRef](#)] [[PubMed](#)]
14. Ghafoor, M.B.; Ujjan, I.D.; Waryah, A.M.; Memon, F.A. Spectrum of HBB gene variants and major endocrine complications in thalassemia patients of Pakistan. *JPMA J. Pak. Med. Assoc.* **2023**, *73*, 2013–2016. [[CrossRef](#)] [[PubMed](#)]
15. Risoluti, R.; Materazzi, S.; Sorrentino, F.; Bozzi, C.; Caprari, P. Update on thalassemia diagnosis: New insights and methods. *Talanta* **2018**, *183*, 216–222. [[CrossRef](#)] [[PubMed](#)]
16. Cappellini, M.D.; Viprakasit, V.; Taher, A.T.; Georgiev, P.; Kuo, K.H.; Coates, T.; Voskaridou, E.; Liew, H.K.; Pazgal-Kobrowski, I.; Forni, G.L.; et al. A phase 3 trial of luspatercept in patients with transfusion-dependent β -thalassemia. *N. Engl. J. Med.* **2020**, *382*, 1219–1231. [[CrossRef](#)] [[PubMed](#)]
17. Thiele, H.; Nürnberg, P. HaploPainter: A tool for drawing pedigrees with complex haplotypes. *Bioinformatics* **2005**, *21*, 1730–1732. [[PubMed](#)]
18. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–424. [[CrossRef](#)] [[PubMed](#)]
19. Yusifova, F.; Akhundova, E.; Salaeva, S.; Gasimov, K.; Shahmuradov, I. Possible relationships between single nucleotide polymorphisms and expression of human HBA1, HBA2, HBB and HBD genes. *Trans. Inst. Mol. Biol. Biotechnol. AR MSE* **2023**, *7*, 50–61.
20. Ayaz, M.; Muzammal, M.; Siraj, S.; Fatima, S.; Fatima, S.; Khan, J.; Khan, M.A.; Shah, M.I.; Rehman, Z.U.; Wei, L. Genetic basis of β -thalassemia in families of pashtun ethnicity in Dera Ismail Khan district of Khyber Pakhtun-Khwa province, Pakistan. *Expert Rev. Hematol.* **2023**, *16*, 693–699. [[CrossRef](#)] [[PubMed](#)]
21. Yasmeen, H.; Toma, S.; Killeen, N.; Hasnain, S.; Foroni, L. The molecular characterization of Beta globin gene in thalassemia patients reveals rare and a novel mutation in Pakistani population. *Eur. J. Med. Genet.* **2016**, *59*, 355–362. [[CrossRef](#)] [[PubMed](#)]
22. Ur Rehman, S.; Khan, J.; Khan, R.M.; Azam, M.; Mutahir, Z. Molecular heterogeneity of β -Thalassemia in Karak District, Khyber Pakhtunkhwa, Pakistan. *Pak. J. Zool.* **2022**, *54*, 45–458. [[CrossRef](#)]
23. Thein, S.L. The molecular basis of β -thalassemia. *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a011700. [[CrossRef](#)] [[PubMed](#)]
24. Rehman, S.U.; Shakeel, M.; Azam, M.; Akhtar, S.; Ziaullah Niazi, R. Frequencies of beta thalassemia mutations show different pattern in Bannu Region than other parts of pashtun population in Khyber Pakhtunkhwa Province Pakistan. *Indian J. Hematol. Blood Transfus.* **2021**, *37*, 479–483. [[CrossRef](#)] [[PubMed](#)]
25. Tanveer, M.I.; Farzand, S.; Manzoor, A.; Anwar, A.; Farhan, S.; Moin, S. Molecular Characterization of Beta Thalassemia Major in Punjab, Pakistan. *J. Ayub Med. Coll. Abbottabad* **2026**, *37*, 334–338.

26. Hussain, A.; Ahmed, S.; Ali, N.; Mailk, H.S.; Anees, M.; Chuahdry, A.H.; Ahmed, P. Rare β -globin gene mutations in Pakistan. *Hemoglobin* **2017**, *41*, 100–103. [[CrossRef](#)] [[PubMed](#)]
27. Sabiha, B.; Haider, S.A.; Jan, H.; Yousafzai, Y.M.; Afridi, O.K.; Khan, A.A.; Ali, J. Development of the next generation sequencing-based diagnostic test for β -thalassemia and its validation in a pashtun family. *Hemoglobin* **2020**, *44*, 254–258. [[CrossRef](#)] [[PubMed](#)]
28. Moassas, F.; Nweder, M.S.; Murad, H. Hb Knossos (HBB: C. 82G> T), β -globin CD 5 (– CT)(HBB: C. 17_18delCT) and δ -globin CD 59 (– a)(HBD: C. 179delA) mutations in a Syrian patient with β -thalassemia intermedia. *BMC Pediatr.* **2019**, *19*, 61. [[CrossRef](#)] [[PubMed](#)]
29. Sadoon, T.A.; Othman, H.E. Complete molecular spectrum of β -globin gene mutations via direct sequencing identifies seven novel variants in β -thalassemia major. *PLoS ONE* **2025**, *20*, e0336610. [[CrossRef](#)] [[PubMed](#)]
30. Farmakis, D.; Porter, J.; Taher, A.; Cappellini, M.D.; Angastiniotis, M.; Eleftheriou, A. 2021 Thalassaemia International Federation guidelines for the management of transfusion-dependent thalassemia. *Hemasphere* **2022**, *6*, e732. [[CrossRef](#)] [[PubMed](#)]
31. Sadiq, I.Z.; Abubakar, F.S.; Usman, H.S.; Abdullahi, A.D.; Ibrahim, B.; Kastayal, B.S.; Ibrahim, M.; Hassan, H.A. Thalassemia: Pathophysiology, diagnosis, and advances in treatment. *Thalass. Rep.* **2024**, *14*, 81–102. [[CrossRef](#)]
32. Huits, R.; Feyens, A.M.; Lonneville, N.; Peyrassol, X.; Adam, A.S.; Gulbis, B.; Van Esbroeck, M. Diagnosis and clinical relevance of co-inheritance of haemoglobin D-Punjab/ β +thalassemia traits in an immigrant Afghan family. *J. Clin. Pathol.* **2022**, *75*, 861–864. [[CrossRef](#)] [[PubMed](#)]
33. Bilgen, T.; Clark, Ö.A.; Öztürk, Z.; Yeşilipek, M.A.; Keser, İ. Gap-PCR Screening for Common Large Deletional Mutations of β -Globin Gene Cluster Revealed a Higher Prevalence of the Turkish Inversion/Deletion ($\delta\beta$)0 Mutation in Antalya. *Turk. J. Haematol.* **2016**, *33*, 107–111. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.