Studies on polymorphism in exon 5 of the *GYPB* gene in malaria positive patients in Punjab, Pakistan

BY

AREEJ MUSHTAQ

Department of Zoology

Faculty of Biological Sciences

Quaid-i-Azam University

Islamabad, Pakistan

2023

Studies on polymorphism in exon 5 of the *GYPB* gene in Malaria-positive patients in Punjab, Pakistan

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Philosophy

In

PARASITOLOGY

By

AREEJ MUSHTAQ

Department of Zoology

Faculty of Biological Sciences

Quaid-i-Azam University

Islamabad, Pakistan

2023

IN THE NAME OF ALLAH

DECLARATION

I hereby declare that the material and information presented in this thesis in my original work. I have not previously presented any part of this work "**Studies on polymorphism in exon 5 of the *GYPB* gene in Malaria positive patients in Punjab, Pakistan.**"

Areej Mushtaq

Dedicated

To my grandfather, Haji Muhammad Ishaq, without whom it was not possible

Table of Content

ACKNOWLEDGMENT

All praise to the cause of causes!

***ALLAH**, the Ultimate Source of creation and goodness. It is due to Him that I found the drive to work diligently, and His guidance smoothed my journey, resolving every concern that came my way.*

*I would like to express my gratitude to my worthy Supervisor, **Dr. Kiran Afshan**, for believing in me, encouraging me, and showing me the path whenever it was dark. I would like to extend my sincere appreciation to Chairperson **Dr. Amina Zuberi** for her unwavering support and assistance in every possible manner.*

*I am highly indebted to my grandfather **Muhammad Ishaq** for his genuine cooperation and help throughout my academic journey. This has left an indelible mark on my heart.*

*A special note of gratitude to my dear mama, **Safia Mushtaq**, whose unwavering belief, in me lights my path, and to my Baba, **Muhammad Mushtaq** and my brother, **Muhammad Hassaan**, who made each and every day of this journey easier. My family's trust, generous financial support, and ceaseless encouragement have been the cornerstones of my academic journey.*

*Especial thanks to my seniors and lab fellows, **Maria Komal**, **Riffat Naseem** ,**Zainab Akhtar**, and **Tayyaba Shan**, for guiding me and assisting me in every possible way.*

*With a truly thankful heart, I want to extend appreciation to my dear friends, **Shehzeen Fatima** and **Noumana Malik**. Without them, this journey might have felt like an uphill climb.*

Areej Mushtaq

List of Tables

List of Figures

List of Abbreviations

ABO	Blood group system (A, B, AB, O)
BLAST	Basic Local Alignment Search Tool
CR1	Complement receptor 1
DNA	Deoxyribonucleic Acid
EBA	Erythrocyte binding antigen
EBAs	Erythrocytes Binding Antigens
GPC	Glycophorin C
GPA	Glycophorin A
GPB	Glycophorin B
GYPE	Glycophorin E
GYPA	Glycophorin A gene
GYPB	Glycophorin B gene
GYPE	Glycophorin E gene
IRBC	Infected red blood cell
MRDTs	Malaria rapid diagnostic tests
PCR	Polymerase Chain Reaction
PK	Proteinase Kinase
RBC	Red blood cell
RBCs	Red Blood Cells
RH	Reticulocyte binding-like homolog
RNA	Ribonucleic acid
SDS	Sodium Dodecyl Sulphate
SNPs	Single Nucleotide Polymorphisms
UAI	Unique Anonymous Identification
UTR	Untranslated Region
WHO	World Health Organization
UCSC	University of California, Santa Cruz
EDTA	Ethylenediaminetetraacetic Acid
TBE	Tris-Boric Acid EDTA

ABSTRACT

Background: Malaria, caused by the *Plasmodium* parasite, continues to be a major worldwide health problem, particularly in endemic countries, such as Pakistan. The purpose of this study is to examine the polymorphism in exon 5 of the *GYPB* gene in individuals who tested positive for malaria in Khushab, Punjab, Pakistan.

Methodology: Cross-sectional study involved the collection of blood samples from malaria-positive patients for the extraction of genomic DNA by the phenol-chloroform method. Primers were designed, optimized, and used for PCR amplification of Exon 5 of GYPB gene. PCR products were confirmed through agarose gel electrophoresis, succeeded by Sanger sequencing and mutation analysis using bioinformatics tools.

Results: Among 35 malaria-positive cases, 52.14 % had *P. vivax* infection , 28.57% had *P. falciparum* infection , and 14.29 % mixed infections. Sequencing revealed polymorphisms and mutations in the exonic, intronic, and 3'UTR regions of the GYPB gene. Three of mutations were already reported and 5 are the novel ones. Samples ME5.32 (c.272C>A) and ME5.18 (g.41355T>C) displayed reported mutation, resulting in an alteration of an amino acid (A91E). Novel mutations were identified in samples ME5.32, ME5.24, ME5.21, characterized by g.144462A>G, g.144507C>T, and g.144498T>C, respectively. ME5.18 demonstrated novel polymorphisms and mutations: g.144522_144523insG, g.144498T>C, and g.41355T>C.

Conclusion: The study revealed unique genetic variations within the GYPB gene, among malaria-infected individuals in Khushab. Pakistan. These findings highlight the need for further widescale investigations to understand the genetic landscape of malaria in the Pakistani population and its significance for disease susceptibility and treatment.

INTRODUCTION

1.1 . Background and Rationale

Malaria stems from the bite of an infected female mosquito of the genus *Anopheles*, transmitting *Plasmodium* parasites to humans and initiating this complex and vital health concern (Talapko *et al*., 2019). Global malaria mortality rates range between 0.3% and 2.2%. with severe malaria cases occurring in tropical regions, ranging from 11 to 30% (White *et al*., 2014). Six species from the genus *Plasmodium—Plasmodium falciparum, P. knowlesi, P. vivax, P. malariae, P. ovale curtisi, and P. ovale wallikeri*—are responsible for transmitting malaria (Garrido *et al*., 2019).

This amoeboid intracellular parasite , *Plasmodium* gathers malaria pigment, an insoluble by-product of hemoglobin (Talapko *et al*., 2019). Most malaria infections are caused by *P.falciparum*, although *P.vivax* are estimated to cause 8% of cases, worldwide (Caputo *et al*., 2016). *P. falciparum* malaria is more dangerous and can frequently result in cerebral malaria and death, especially in youngsters. When compared to *P. falciparum* infection, malaria caused by *P. vivax* was first thought to be milder and more treatable, but new studies indicate that *P. vivax* may result in complications that lead to death. Recorded death rate for *P. vivax* is between 0.1% and 1.6% , worldwide (Prince *et al*., 2009).

The total population of *Plasmodium* hosts stages from one to more than 106 organisms, replicates in ten or more morphological stages, and can be as little as one cell or as many as 10,000+ cells. A limited number of these morphological phases induce clinical disease in the human host and people with malaria around the world exhibit few symptoms (Danny *et al*., 2018). Malaria is a significant contributor to early pregnancy loss, low birth weight, severe illness, and infant mortality in both pregnant women and young children (Caputo *et al*., 2016). Plasmodium falciparum infection continues to present a significant public health concern on a global level, as it is responsible for over 90% of malaria-related fatalities(Zekar *et al*., 2020).

Plasmodium spp. have a complex life cycle that alternates between female Anopheles mosquitoes and vertebrate hosts and necessitates the creation of special zygote forms to infect various cell types at particular phases. The asexual cycle in the blood begins when sporozoites

have infected the host's hepatocytes. A mosquito that is feeding consumes the sexual forms that emerge during the blood stage to complete the cycle (Cowman *et al*., 2016).

The development of medication resistance in *Plasmodium* species is a significant barrier to the control of the disease, even though malaria is curable. Understanding the molecular basis of the pathophysiology of malaria is essential , to facilitating the development of novel techniques to combat the disease due to insecticide resistance in disease-transmitting mosquitoes (Tuteja *et al* ., 2007)

1.2 . Taxonomic Categorization of *Plasmodium* Species

Kingdom: Chromista

Phylum: Apicomplexa

Class: Aconoidasida

Subclass: Coccidia

Order: Haemosporida

Family: Plasmodiidae

Genus: *Plasmodium*

Species: *P. falciparum, P. knowlesi, P. vivax, P. malariae*

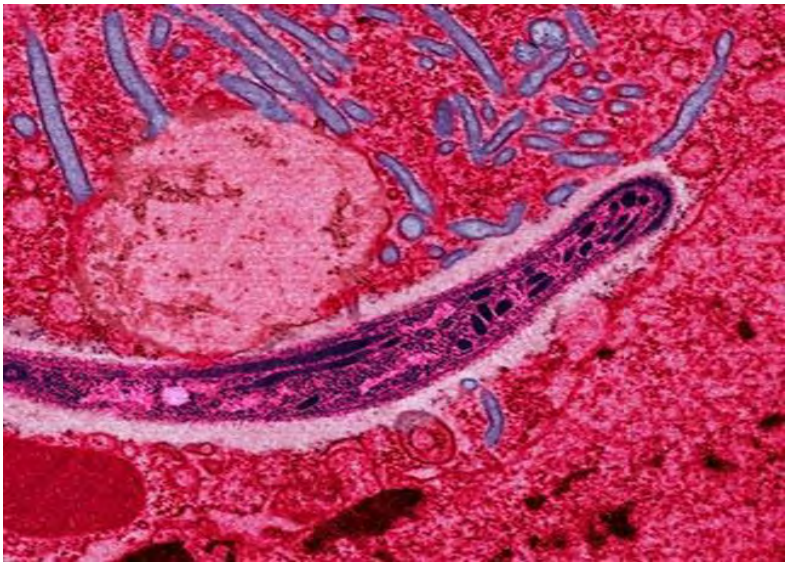

Fig. 1.1: Electro micrograph of a sporozoite stage of *Plasmodium*

1.3 . Life Cycle of *Plasmodium*

Asexual cycle occurs in man, the vertebrate host, and the sexual cycle occurs in the vector Anopheles, the invertebrate host, as part of *Plasmodium's* life cycle. Sporozoites, an infectious type of *Plasmodium*, are transmitted when a female Anopheles mosquito bites a human host (Basu *et al*., 2017). The sporozoites enter the bloodstream, travel to the liver, exit the sinusoids through Kupffer or endothelial cells, and then enter a hepatocyte. Cellular exploration occurs before active invasion until a suitable hepatocyte is located.

Tens of thousands of merozoites are discharged into the vasculature in packets of merosomes after they form a PVM and go through schizogony. They come into contact with erythrocytes there, starting a long-lasting cycle of asexual schizogony in the bloodstream. A portion of merozoites that reproduce asexually are reprogrammed to go through gametocytogenesis.

Gametocytes sequester and develop in the bone marrow within 15 days. Once they are fully developed, they exit from the bone marrow and enter the peripheral blood to be consumed by a mosquito. There, they emerge as extracellular male and female gametes in the midgut. Micro- and macrogametes combine to form a zygote, which develops in 24 hours into an ookinete and migrates through the mosquito's midgut epithelium before encysting to become an oocyst, where asexual sporogonic reproduction takes place. Oocyst rupture releases mobile sporozoites into the hemocoel, which then travel to the salivary glands from where they can be injected into the human host with the bite of female Anopheles mosquito (Cowman *et al*., 2016).

Interactions between the host, the vector, and the parasite are necessary for the transmission of malaria. The parasite cannot finish development process outside the range of specific temperatures that varies based on different *Plasmodium* species. The parasite completes the sporogonic cycle inside the mosquito and the asexual cycle in the human host. To survive the winter in cold areas, *P. vivax* has developed a variety of coping mechanisms. The life cycle of *Plasmodium falciparum* is shown in Fig.1.2.

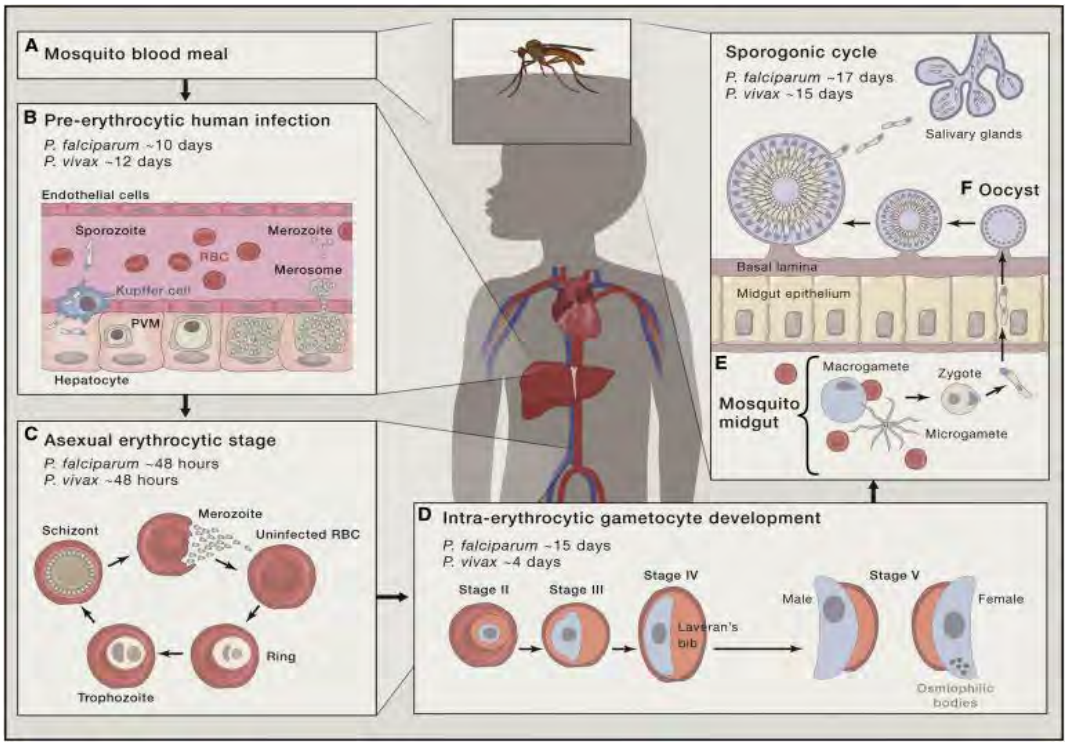

Fig. 1.2: Life Cycle of *Plasmodium Falciparum*

The elements needed by the parasite to complete its life cycle are the ambient temperature, the existence of vectors appropriate for the transmission, and susceptible hosts. Instead, more than one factor, not just the temperature outside, is needed to support mosquito survival in the environment. In the case of vector-borne diseases, humans are the third actor and only experience the indirect effects of climate change. Extreme environmental conditions, like floods and drought, can cause people to move, but the disease won't change in terms of transmission without parasites and vectors (Caputo *et al*., 2016).

1.4 . Malaria Pathogenesis

Multiple clinical signs and symptoms of severe malaria can appear in several anatomical locations. The pathogenicity of the severe forms of the disease is influenced by factors that are relevant to both the parasite and the host. The capacity of parasites to adhere to the vascular endothelium is known as cytoadherence. Mature forms of the parasite stage can attach to the placenta, subcutaneous adipose tissues, and the vascular endothelium of many organs (Autino *et al*., 2012). The pathophysiological basis of severe manifestations of malaria, such as cerebral malaria, is believed to be parasite sequestration

(Grau *et al*.,2012). It impairs blood flow, which produces localized hypoxia. It speeds up parasite reproduction and IRBC adhesion to uninfected red blood cells (Dondorp *et al* ., 2004).

About 20 hours after the RBC invasion, mature parasite forms play a major role in sequestration. The parasites generate fresh proteins that are transferred to the surface of IRBCs, making them more adherent to the endothelium. The parasites can hide out for 24 hours during their 48-hour life cycle in the deep microvasculature. Because they are not visible in the peripheral blood, they avoid splenic clearance in this way, complicating the diagnosis (Autino *et al* ., 2012). Scanning electron micrograph of *Plasmodium*-infected red blood cells is shown in Fig.1.3.

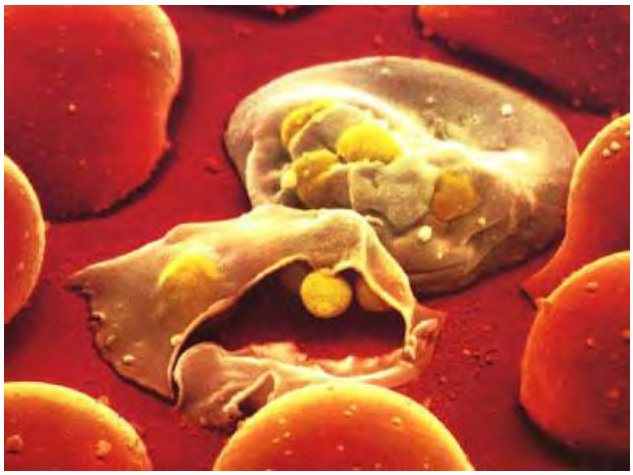

Fig. 1.3: Scanning Electron micrograph of *Plasmodium*-infected red blood cell

Rosetting is one of the cytoadherence mechanisms of late-stage, IRBC to non-parasitized red blood cells and/or platelets (Rowe *et al*., 2009). *P. falciparum, P. vivax* and *P. ovale* can all produce rosettes, but only *P. falciparum*-caused rosettes have been linked to severe malaria (Doumbo *et al.,* 2009).

In essence, the immune response to the parasite is species- and stage-specific, and it is complex and not well understood (Stevenson *et al*., 2004). To prevent parasite proliferation and ultimately help with infection removal and resolution, the innate immune system's components must be activated (Urban *et al*., 2005). The innate immune system's cells, including neutrophils, monocytes/macrophages, dendritic cells, natural killer (NK) cells, NKT cells, and gamma T cells, are incharge of stopping the early stages of the disease through phagocytosis and/or the creation of inflammatory mediators.

The inflammatory response organized by the cells of the innate immune system, driven by parasites or their products after the rupture of the late-stage infected erythrocytes, is what causes many of the symptoms of malaria attacks, including fever, nausea, headaches, and others (Taramelli *et al* ., 2000)

A majority of the proteins exported to the surface of the *P. falciparum*-infected erythrocyte have been characterized by members of the PfEMP1 family. The parasite can switch expression of antigenically different isoforms, a process known as antigenic variation, to elude immune destruction as a result of their display on the erythrocyte surface, which exposes them to host antibodies. The great diversity and tightly controlled expression of PfEMP1 proteins play a significant role in *P. falciparum* ability to sustain a protracted infection. They enable parasites to hide out in deep capillary beds and avoid splenic clearance thanks to their diverse array of sticky domains (Nunes-Silva *et al*., 2015).

1.5 . Malaria Incidence and Distribution

Global Prevalence of Malaria

According to World Health Organization (WHO) many studies conducted, between 2000 and 2021, there were an estimated 2 billion cases of malaria worldwide, preventing 11.7 million deaths, (World Health Organization , 2022). Populations that are susceptible to malaria vary in endemic areas depending on the rate of transmission and level of immune development. According to the World Health Organization-recommended new malaria control measures, children under five years and expectant women are its key targets in highly endemic areas (World Health Organization , 2010). A rise from 245 million cases in 2020 to a projected 247 million cases worldwide in 84 malaria-endemic countries in 2023, with the majority of this increase coming from nations in the WHO African Region (World Health Organization, 2022).

Malaria Incidence in Pakistan

Pakistan suffers from a malaria epidemic. As opposed to the 2.6 million suspected cases recorded in 2021, more than 3.4 million cases of malaria were reported across Pakistan from January through August of 2022. More than 170,000 cases that were lab-confirmed were *Plasmodium vivax*. Following the severe floods in mid-June 2022, a sharp increase in cases was seen in the provinces of Balochistan and Sindh, which together accounted for

78% of the total confirmed case (World Health Organization, 2022). Fig1.4. highlights malaria-endemic areas in Pakistan.

Fig.1.4: Malaria-affected regions of Pakistan in 2022

Pakistan continues to carry a sizable amount of the burden in the fight against global malaria, with an estimated 1 million cases annually. *P. falciparum* and mixed cases (double infections with *P. vivax and P. falciparum*) account for 14.9% and 1.1% of cases, respectively, while *P. vivax* malaria is currently 84% frequent in Pakistan, according to the WHO. Tribal regions (formerly known as FATA: Federally Administered Tribal Places) in Khyber Pakhtunkhwa (KPK), which are the poorest and most severely underdeveloped areas of Pakistan, have the highest malaria load due to the high number of Afghan refugees and internally displaced people (IDPs) (Karim *et al*., 2021)

1.6. Diagnostic Approaches for Malaria

With around 30,000 annual deaths, malaria is the second most common disease in Pakistan. To reduce the chances of deadly consequences linked with malaria in endemic countries like Pakistan, accurate and prompt diagnosis is essential (Mukry *et al* ., 2017). Finding malaria parasites, antigens, or products in the patient's blood is necessary for malaria

diagnosis. Various manifestations of the five different types of malaria; the various stages of erythrocytic schizogony; the endemicity of various species; the relationships between the degree of transmission, population distribution, parasitemia, their immunity, and signs and symptoms; drug resistance; the issues with recurrent malaria; persisting viable or non-viable parasitemia; and sequestration of the parasites in the deeper tissues; and the effect of chemotherapy treatment or even presumptive treatment based on clinical diagnosis, can all influence the identification and interpretation of malaria parasitemia in a diagnostic test (Tangpukdee *et al*., 2009)

Clinical assessments are determined by the patient's indications and symptoms in addition to physical examination results. Early malaria symptoms vary widely and are fairly non-specific, including fever, weakness, myalgia, chills, sense of disorientation discomfort in the stomach, diarrhea, nausea, and vomiting together with anorexia and pruritus (Looareesuwan *et al*., 1994). Malaria symptoms are often accompanied by substantial RBC lysis at the blood stage throughout the lifecycle (Molina-Franky *et al*., 2022). Due to the non-specific character of the signs and symptoms, which have a great deal of overlap with other frequent and potentially fatal diseases, such as frequent viral or bacterial infections, and other severe illnesses, a clinical diagnosis of malaria is still difficult (McMorrow *et al*., 2008).

At present, a few techniques are available for diagnosing malaria. The use of light microscopy to evaluate stained peripheral blood smears is one of the traditional procedures for clinical diagnosis, along with history-taking and physical assessment, empirical/syndromic diagnosis (mostly the presence of fever in endemic regions), and light microscopy. Histopathology has a limited function, although despite being helpful in specific circumstances, it is not helpful in malaria control initiatives. The diagnosis of malaria is seldom ever based on nucleic acid amplification tests, which are only used in a few public health laboratories and are not widely available. As with other prevalent infectious illnesses, several fast diagnostic tests have been created and sold. They are known as MRDTs, and they could affect malaria diagnostic and treatment programs the most (Wilson *et al*., 2013). Due to its accessibility and low cost, the microscopic identification of malarial parasites is typically regarded as the gold standard in the diagnosis of malaria. Although inexpensive, precise, and sensitive, this method needs a skilled

microscopist and may be unreliable at parasite counts of less than 1000 parasites per liter (Mukry *et al*., 2017).

1.7. Vectors in Transmission of malarial parasites

The Culicidae family of mosquitoes (Diptera) is the major group of disease-carrying insects. Mosquitoes pass their larval stage in water and adult stage on land . The Anopheles mosquito is the carrier of the mammalian parasite *Plasmodium sp*., which causes malaria. About 50, of the approximately 500 species, of Anopheles are capable of transmitting malaria to humans (Harbach *et al*., 1994).

In real terms, the majority of worldwide transmission is caused by 20 Anopheles species. The other mosquitoes don't contribute to the transmission because they either selectively bite the animal or are immune to *Plasmodium* or a strain of *Plasmodium*. There are variations in the ability of distinct Anopheles species to transmit different *Plasmodium* types. Certain *Anopheles* and *Plasmodium* species' capacity to transmit varies regularly and is genetically determined depending on the geographic origin. The major vectors, which have a wide geographic distribution, the secondary vectors, which have a more focused relevance, and the undefined secondary vectors are distinguished based on the geographic areas (Mouchet *et al* ., 2004).Fig.1.5 represents the vector of *Plasmodium sp, Anopheles gambiae*.

Fig.1.5 : *Anopheles gambiae* vector for *Plasmodium sp.*

1.8. *Plasmodium* Parasite Receptor interactions

Plasmodium falciparum, the deadliest malaria parasite in humans, employs diverse invasion pathways to engage with host cell surface receptors (Molina-Franky *et al*., 2022). It modifies the structure of infected red blood cells, increasing their stiffness, adhesiveness, and susceptibility (Molina-Franky *et al*., 2022). Given that mammalian erythrocytes lack nuclei and biosynthetic pathways, *P. falciparum* primarily infects such cells (Mei *et al*., 2021). *P. falciparum* may exploit polymorphism and antigenic variation to elude the immune system of humans. These invasion pathways involve redundant ligands binding with host-specific receptors to penetrate RBCs (Gilson *et al*., 2017). Fig.1.6 highlights binding sites of *Plasmodium* protein ligands to receptors on RBCs.

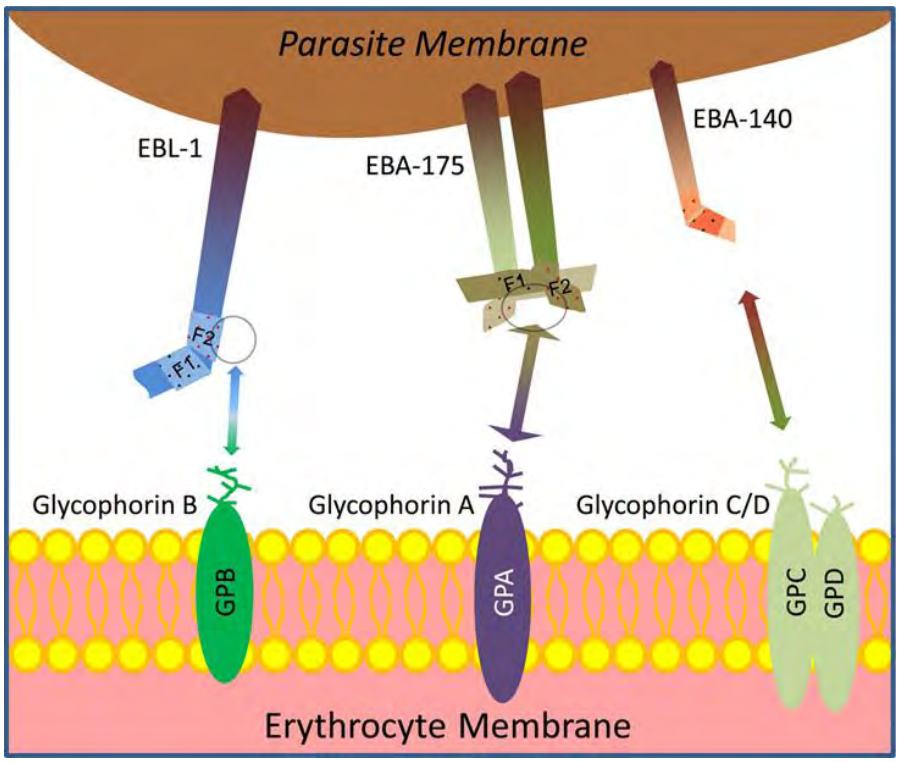

Fig.1.6: Binding sites of *Plasmodium* protein ligands to receptors on RBCs

For successful invasion of human red blood cells (RBCs) by *Plasmodium falciparum* merozoites, there must be precise interactions between the parasite's ligands and the host RBCs' surface receptors. The host RBCs contain receptors such as glycophorin A (GPA), glycophorin B (GPB), basigin (CD147), and complement receptor 1 (CR1), whereas prominent ligands encompass erythrocyte binding antigens (EBAs), reticulocyte binding-like homologs (RHs), with merozoite surface proteins (MSPs). Understanding the molecular basis of these

ligand-receptor interactions is critical for devising focused anti-malarial therapy approaches and it also enhances our understanding of host specificity during malaria invasion (McGhee *et al*., 1953).

During the invasion of human red blood cells, *Plasmodium falciparum's* EBL ligand interacts with erythrocyte glycophorin receptors (Li *et al*., 2012). ECL-1 plays a role in merozoite invasion by interacting with the erythrocyte glycophorin receptor GPB. The human erythrocyte receptors for the *P. falciparum* ligands EBA-175 and EBA-140, respectively, have been identified as GPA and GPC/D (Li *et al*., 2012).

The discovery of EBL-1, the second member of the EBL family, was guided by consensus sequence homology (Peterson *et al*., 1995). EBL-1 undergoes missense mutations, leading to a shortened protein during the schizont late stage (Peterson *et al*., 2000). According to a theory the selection pressure to block *P. falciparum's* primary invasion pathway may have led to the evolution of GPB (Salinas *et al* ., 2014).

Plasmodium falciparum's preference for human red blood cells (RBCs) following invasion is driven by multiple dynamic ligand-receptor interactions. Through these interactions, the parasite improves its survival and evades immune responses by adapting to changes in host RBCs. *P. falciparum* merozoites utilize various RBC receptors, including Band 3, Glycophorin A (GPA), Glycophorin B (GPB), and Glycophorin C (GPC). Genetic differences in GYPA and GYPB determine the MN and Ss blood types (Cowmen *et al*., 2006) .

1.9 Blood Group System and GYPB Gene

Landsteiner along with his colleagues after the first discovery of the ABO blood group in 1900, discovered another blood group, viz., MNS, in 1927. More than 40 antigens are now included in this category, with M, N, S and s perhaps the most common. Glycophorins, which lie in the RBC membrane, are sugar-bearing transmembrane proteins that carry the antigens of the MNS blood group system. These proteins are enriched with carbohydrates primarily in the form of Sialic acid. One side of the glycophorin protein is anchored to the sugars that define the MNS blood type of a person, and the other extremity is attached to the underlying cell. The MNS antigens are encoded by two genes GYPA and GYPB, located on chromosome 4(4q28.2-q13.1). M and N are codominant alleles found in GYPA, arising from three single nucleotide polymorphisms (SNPs) in the gene, which results in a distinction between the M and N antigens by two amino acids. GYPB, S and s are the codominant alleles that arise from one SNP in the

gene leading to the differentiation of S and s antigens by a single amino acid (Dean *et al*., 2005). The S and s antigens are represented by different phenotypes and are identified by an amino acid shift at position 29 [Met(S)/Thr(s)] of GPB(Wang *et al*., 2003). Fig.1.7 highlights the Amino acid Sequence of GPB indicating the transmembrane domain of protein encoded by exon 5.

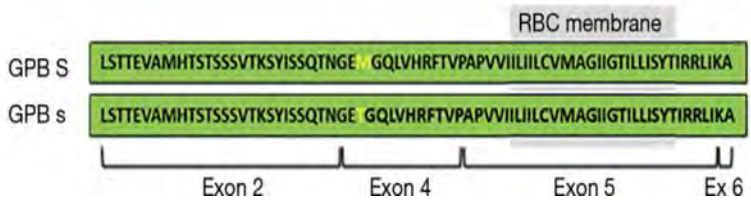

Fig.1.7: Amino acid Sequence of GPB indicating Transmembrane Domain of protein

On chromosome 4 the glycophorins genes i.e., GYPA, GYPB, and GYPE are homologous and contiguous. GYPA contains seven exons, GYPB has six exons (exon 3 is a pseudo exon or non-coding exon), and GYPE has six exons (exons 3 and 4 are pseudo exons). The leader sequence for each of these GYP genes is encoded by exons 1 and 2, the extracellular domains are encoded by exons 2-4, and the transmembrane domains are encoded by exon 5 (Reid *et al*., 1994).

GYPA and GYPB genes are responsible for synthesizing two distinct proteins, GPA which comprises of 150 amino acids, and GPB which is made up of 91 amino acids (Daniels *et al*., 2013). GPA and GBP are transmembrane proteins that are located on the outer layer of the RBCs and have sugar moieties attached to them, particularly O-glycans. Also, it's important to note that N-glycans are only present in GPA and not in GBP. These molecules on the surface of RBCs create an overall negative charge that prevents blood from aggregating inside the blood vessels and maintains the flow of blood in the circulatory system (Huang *et al*., 1995). Moreover, they also act as receptors for various cytokines and pathogens, including the malaria-causing parasite, *P.falciparum.* Fig.1.8 represents the extracellular domain of the GYPB protein,

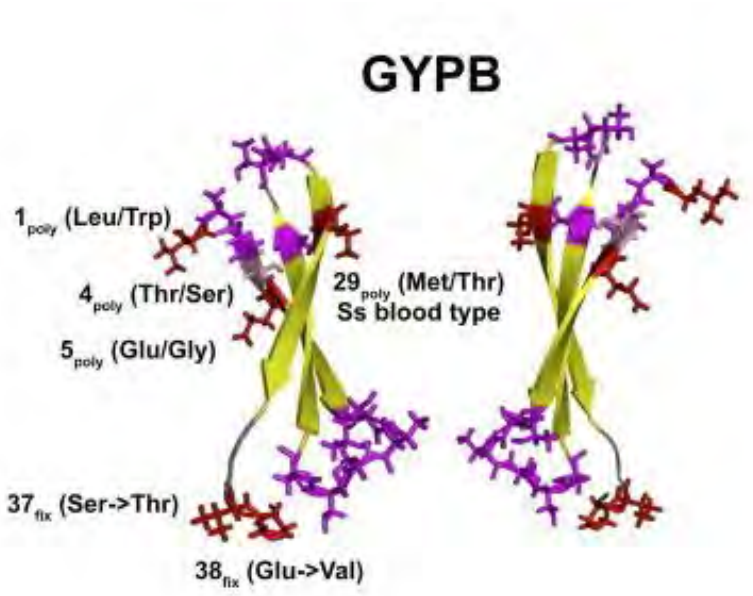

Fig.1.8: Extracellular domain of GYPB

1.10 Impact of GYPB Exon 5 Polymorphism on Malaria Susceptibility

Recent research has shown that the susceptibility or resistance of erythrocytes to *P. falciparum* invasion is greatly influenced by genetic variations in glycophorins and other well-known receptors. The polymorphisms in *P. falciparum* GPA and GPB glycophorin receptors so far have been demonstrated to be only sporadically related to the efficiency of *P. falciparum* invasion and to give only some degree of protection against RBC invasion (Tarazona-Santos *et al.,* 2011).

GYPB nucleotide alterations affecting S or s antigens are expressed on RBCs. The amino acids 33 to 39 characterize the GPB U antigen. As a result, loss of GYPB exons 2–5 leads to S-s-U- phenotype, or lack of GPB on RBCs (Tarazona-Santos *et al*., 2011). Red cells with S-s-U negative phenotype have low levels of glycophorin B and are fairly susceptible to invasion. Different phenotypes are associated with variant GPB which encodes He antigen and its molecular origin is nucleotide changes that occur in or around the GYPB gene's exon 5. Understanding how these genetic differences affect the invasion process will help us to better understand the processes behind host-parasite interactions and help us locate prospective malaria therapeutic targets (Pasvol *et al*., 1982).

Specific haplotypes at or inside the glycophorin gene cluster on human chromosome 4 were shown to be strongly associated with a 33% lower risk of contracting severe malaria (Wassmer *et al* ., 2016). In this study, we are looking for any polymorphism or mutations, in the endemic population of Khushab, that may render glycophorin B fairly susceptible to malarial invasion and lower the risk of severe malaria as well.

1.11 Objectives

- To Study polymorphism in exon 5 of the GYPB gene in malaria-positive patients in Punjab, Pakistan.

MATERIALS AND METHODS

2.1 Ethical Considerations

The present study received ethical approval from the Bioethical Review Committee of the Department of Zoology, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

2.2 Blood Collection

Blood samples were collected from 35 malaria positive patients admitted with different hospitals of District Khushab , Pakistan , having malaria testing facilities .

A sterile syringe was used to draw 5 mL of venous blood from the participant. A Unique Anonymous Identification (UAI) number and the participant's name were written on the label of each EDTA vacutainer tube used to collect blood samples. To avoid clotting, the tubes were gently mixed and kept in the freezer at -4 degrees Celsius.

2.3 Extraction of Genomic DNA

Phenol-chloroform technique, an established organic protocol spread over three days, was used to extract genomic DNA from the blood sample.

Protocol: Phenol-Chloroform DNA Extraction Method

Day 1: The frozen EDTA tubes containing blood samples were allowed to thaw at room temperature for some 10-15 minutes. A Volume of 750 µl of blood from each sample was carefully transferred into the designated Eppendorf tubes. We added 750 µl of sol A to each Eppendorf tube containing the blood samples. Sol A serves as a critical reagent for cell lysis and DNA stabilization. The tubes were kept at room temperature for 20 minutes to facilitate optimal DNA extraction. After the incubation, the Eppendorf tubes were centrifuged at 13,000 rpm for 15 minutes to separate cellular debris, proteins, and contaminants from the DNA-containing supernatant. Half of the supernatant was discarded ensuring the DNA pellet remains undisturbed. A volume of 400 µl of sol A was added to the remaining DNA pellet for further washing and purification. Complete mixing of sol A with the pellet was ensured by vertexing or tapping the tubes. The centrifugation process was repeated to pelletize the DNA when it was

washed with sol A three to four times until a clear DNA pellet was obtained. Once a clear DNA pellet was obtained, the following chemicals were added to each Eppendorf tube: Solution B (400 µl), sodium dodecyl sulfate (SDS) (25 µl), and proteinase kinase (PK) (5 µl). The contents were thoroughly mixed by vertexing to facilitate DNA precipitation and removal of protein contaminants. The Eppendorf tubes were kept at 37°C overnight, allowing for complete digestion of proteins and enzymatic reactions, thereby promoting optimal DNA yield and quality.

Day 2: The incubated Eppendorf tubes were removed at room temperature for a few minutes to create a controlled environment. Fresh Solution D and Solution C+D were prepared on the same day to ensure accurate DNA extraction results. 500 µl of Solution C+D were added to each Eppendorf tube containing the incubated samples to enable selective separation of DNA from other cellular components. The tubes were centrifuged at 13,000 rpm for 15 minutes to pelletize the DNA and. Eppendorf tubes were labelled the upper DNA layer after centrifugation, ensuring proper sample identification. Carefully transfer 500 µl of Solution D to the newly labeled Eppendorf tubes containing the clear DNA layer. Repeat the centrifugation step to pelletize the DNA in the new tubes. Label new Eppendorf tubes to collect the newly separated DNA layer. Isopropyl alcohol (500 µl) and Sodium acetyl (60 µl) was added to each labeled tube to further facilitate DNA precipitation and purification. Centrifuge the tubes to separate the DNA pellet from the liquid phase, ensuring the removal of impurities. Carefully discard the liquid phase while preserving the DNA pellet. Wash the DNA pellet by adding 200 µl of 70% ethanol to each tube, followed by centrifugation to remove residual impurities and ethanol. Allow the tubes to air dry completely to ensure the removal of all ethanol and the absence of visible bubbles. In the final step of Day 2, add 200 µl of TE buffer to each dried tube. Incubate the tubes overnight at 37°C, allowing the DNA to rehydrate and dissolve in the buffer.

A heat shock treatment was to the extracted DNA by placing the tubes in a water bath at 70°C for 1 hour. This step inactivates nucleases and other enzymes, ensuring DNA integrity and preventing denaturation. After the heat shock treatment, the tubes were allowed to equilibrate at room temperature for 5 minutes to minimize thermal stress on the DNA. The tubes were given a short spin by centrifuging them at 3000 rpm for 2 minutes to ensure thorough mixing of the DNA and buffer. Store the DNA samples at -20°C in properly labeled cryo boxes for long-term stability and future analysis. Table 1 in Annexure I enlists chemicals conc. and composition used in DNA extraction.

2.4 Agarose Gel Electrophoresis (1%)

Following DNA extraction, agarose gel electrophoresis is carried out as a verification step. The following is a description of the procedure's protocol:

In a conical flask, combine 50 mL of 1X TBE buffer with 0.5 g of agarose powder to create a 1% agarose gel. To make 1X TBE buffer, dilute 10X TBE buffer in a 1000 mL bottle with 900 mL of distilled water. To generate a final volume of 1000 mL for the 10X TBE buffer, combine 0.5 M EDTA (40 mL), 108 g of Tris, 54 g of Boric acid, and distilled water. Before adding the formed gel to the gel tank, adjust pH to 8 Buffer. To create a clear solution, place the conical flask in the microwave for two to three minutes while it is covered in aluminum foil. A few minutes are spent letting the flask cool off at room temperature. In the flask, Ethidium Bromide is introduced in v. 300 microliters. This intercalating substance is used to identify DNA under UV light; it should be handled carefully because it is carcinogenic. In a gel mold, a casting tray and combs are placed. The casting tray is filled with a clear solution that is poured without creating any bubbles. The polymerization process takes place at room temperature for 30 to 40 minutes. The gel tank is filled with running 1XTBe buffer before the gel is added. The gel is carefully poured into the gel tank after solidification and the combs are gently withdrawn. Three microliters of extracted DNA are combined with three microliters of the loading dye, 6X Bromophenol blue, before placing each sample into a well. The Gel electrophoresis device is closed and set to run for 25 minutes at 120 volts. The Gel is thoroughly examined under UV light when the running is complete using the Gel Documentation System Table 2 in Annexure I enlists the Composition of Agarose gel and other required chemicals.

2.5 FASTA Sequence of Selected Exon

TGTCACCCAGGCTCGAGTGCAAAGTGGCGCAATCTCTGCTCACTGCAGCCTCTGCCTCCC
TGGTTCAAGTGATTCTCGTGCTTCAGCCTCCACAGTAATTGGGACTACAGGTGCGGGCTA
CCATGCCTGACTAATTTTTGTATTTTTTTAATAGCAGAGATGGGGTTTCGCTGTGTTGGCC
AGGCTGGTTTCAAACTCCTGACCTCAAGTGATCCACCTGCCTTGGCCTCCAAAGTGCTGG
GATTACAGGTGTGAGCCACTGCGCCCGGCCCTAATTAGGGTTTTTATAAAACCAAAAGA
ACTTGGCAACACCCCTAGGTACCCTTTAGAAGCCTCCAATTGGCTACAGCCTGTGAAGGA
TTGGCCTGTGACCAATCAGAGGCTGAAGTGGAGTCTTAGCTCATGGTCAAGCAGAGGCT
GAAGTGGAAACTTCTTGTCTTTTTATCACAG**GCATGAGGATGTGGCCTGCATGCTGCC**
TGATCTTGCCTAGAACCGGCTGCACCTGCTGTTCTCTTGTTTATGCAAACTGGCTGC

ACCTGCTATTCCTTTGCTTATGCCCCTACCCCTGGCTATCCTAATTCCCTGTTCTCCT
GCCTCACTATTACTGTATTCTCTACTTCTAAATAAAAATAAAACAAAATACAAACCG
TTTTCATGTATTATAGTCATGATTTGTTATTTCAAGCTATGTTTAAAATGTAAAGCCCATT
ACAGTAGGATAGATGAGAAGAAAGAAGGGTCTGAACAAAGCCAACAGCGGAGAACATG
CAGAAGGGGAATGAAGACTTCAGCCTGGACTCTCATTAGGTGGTTGTCCAAGATAACAT
GAAGTCTCATGTGAAATATATTGGAAAAGAACAAGATTCAGTCAAGTACCCAAGGTGCC
TGACTGCTGAAACGGCACTGTAATTTAAAAAGCAAAGACAAGTTTATCATACTGTGAGA
TCCTTGAGGGCAAGGACCAAATTCTATTCAAATCTTTGTCAAAAACCAGATTTTTGATAG
GCTTACTTTAGTATCCATTGATCTTAATTATAGTTCACCTGCCAGGAGTTAAGGTCCCTCC
TCAACAGTAAAGAGAATGACAGGCACAGGGGTAGGAGCACGTAGAGAAGGGCCTGACT
TTCAAATGAACTCTGGAGAGCACTCATCACCTCCAGAACACAGCAGATCTTACTGAGATG
GCTGGGTCAGAAATCTCTGAGGGACTCCAGGATGCCCACGCCTCAGAAAGGACAGGTTA
AACCAAAG

2.6. Primer Designing

Primers were designed to amplify Exon 5 of the Glycophorin B gene on chromosome 4 using the Primer-3 software (https://primer3.ut.ee/). The design process involved optimizing the annealing temperature of the primers, amplicon size, salt concentration, and primer length to ensure their efficacy. The reference sequence necessary for primer design was obtained from the Ensemble website (https://asia.ensembl.org/Homo sapiens/Info/Index). To confirm the specificity of the selected primers, the Blast-like alignment tool (BLAT) on the UCSC genome browser (https://genome.ucsc.edu/cgi-bin/hgBlat) was utilized. Additionally, the In-silico PCR tool on the UCSC genome browser (https://genome.ucsc.edu/cgi-bin/hgPer) was used to verify the amplicon size for the primers. Detailed information, including the locus of mutation, primer sequences, product size, and melting temperature for each primer, is provided in Table 2.3.

Table 2.1. Primer Information for Amplification of Exon 5 in GYPB Gene (Chrom. 4)

Gene	Exon	Locus	Primer Type	Primer Sequence (5'-3')	Primer Length (bp)	Product Size (bp)	Melting Temp. (°C)
GYPB	5	Chrom 4	Forward	GGCTGAAGTGGAGTCTTAGC	21	373	54
			Reverse	TGGCTTTGTTCAGACCCTTC	20		52

2.7. Primer Dilution

The initial concentration of primers ordered was 100 picomole/µl. Further dilutions were prepared by adding PCR water, and final concentrations of 10 picomole/µl were made.

2.8. Primer Optimization

To optimize primers, Gradient PCR was used with 55° optimal annealing temperature for the primer pair, used to analyze the polymorphism within exon 5 of the GYPB gene.

2.9. Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) was used to amplify genomic DNA samples from patients who tested positive for malaria. PCR tubes (Axygen USA) with a 200 L capacity were used for the PCR reactions. The chemical concentration and volume used in the reaction mixture are listed in Table 3 in Annexure I.

The PCR tubes were given a spin in a microfuge at 3000 rpm for 1 minute to ensure thorough mixing before being placed in the thermocycler for the PCR reaction. Table 2.2 lists the conditions of PCR cycles with Fig.2.1. also showing PCR conditions.

Table 2.2: Conditions of PCR Cycles

Step	Temperature	Time	Cycle
Initial Denaturation	96°C	5 min	1X
Denaturation	95°C	45 sec	40X
Annealing	65°C	45 sec	40X
Extension	72°C	60 sec	40X
Final Extension	72°C	10 min	1X
Hold	25°C		

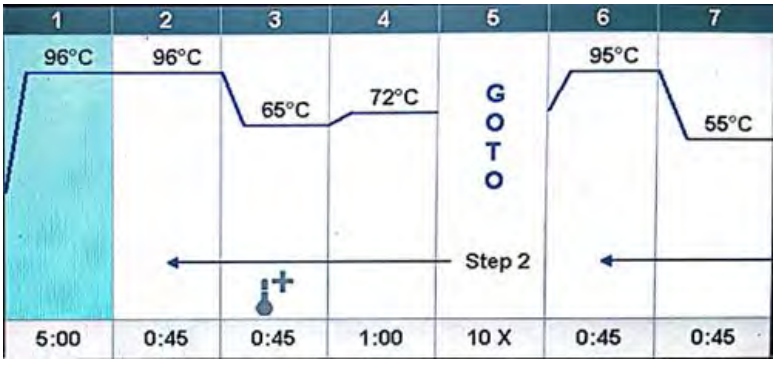

Fig.2.1: Conditions of PCR Cycles

2.10. PCR Product Confirmation

A 2% agarose gel was made to validate the PCR results. To do this, 1g of agarose powder was mixed with 50 ml of 1X T.E. buffer and 2ul of ethidium bromide. The next step was to combine 3µl of 6X fluorescent dye (bromophenol blue) with 3µl of each sample (PCR product). These DNA samples were electrophoresed on a 2% agarose gel for 40 minutes at 120V. Gel Documentation System was then used to analyze the gel to verify the amplification of the desired GYPB Gene segment.

2.10. Sanger Sequencing

The processed PCR samples were transferred for commercial sequencing to Macrogen. Korea. At this center, Sanger's sequencing was conducted, using the big dye terminator chemistry method with an automated ABI PRISM® 3730 Genetic Analyzer. The labeled DNA fragments were separated using capillary electrophoresis, and their detection was confirmed through spectrum analysis. Each nucleotide (A, T, C, and G) was color-coded for documentation purposes. For visualization of Sanger's sequencing outcomes, the software tools Sequencher 5.4.6 and Chromas 2.6.6 were employed.

2.11. Mutation Analysis

The sequence data of samples was analyzed by using Sequencher 5.4.6 software. Polymorphism prediction for each variant was done by various bioinformatics software and tools including MutationTaster (http://www.mutationtaster.org/), polyphen-2 (http://genetics.bwh.harvard.edu/pph2/),and public database frequency was also determined. Public databases including ExAC (http://exac.broadinstitute.org/), 1000 Genomes Browser (http://browser.1000genomes.org/index.html) and SNPs (http://www.ncbi.nlm.nih.gov/SNP/

) Were consulted to determine if the variants were rare.

2.12 Analysis of Demographic Data

The percentage analysis was done using Microsoft Excel and bar charts were constructed for the representations of data.

RESULTS

3.1. Baseline Characteristics of Study

The demographic characteristics of the blood sample donors was collected represent a diverse malaria-affected population in Khushab, Pakistan. The blood sample of 5 malaria-positive patients' blood gathered from different hospitals across the city. Among the participants, various types of malaria infections were diagnosed by Malaria Rapid Diagnostic Tests (RDTs) and Malaria Parasite Microscopy. The infections included cases of *Plasmodium falciparum* infections, *Plasmodium vivax* infections, and mixed infections of both species. Basic data such as names, genders, ages, types of infections, and blood type of patients were collected. The affected patients had general symptoms of malaria including fever, chills, joint pain, and abdominal discomfort.

3.2.1 Malarial cases according to the type of infection

Among 35 individuals with malaria infection, 20 patients had infection due to *P. vivax*, accounting for 52.147 % of cases. 10 patients had an infection due to *P. falciparum* constituting 28.57% of cases, and mixed infections were seen in 14.29 % of cases, which is represented in Table 3.1. and Fig 3.1.

Table 3.1: Percentage of malaria-positive patients against the type of infection.

Infection Type	No. of Patients	Percentage
P. vivax	20	57.1%
P. falciparum	10	28.5%
Mixed	5	14.3%

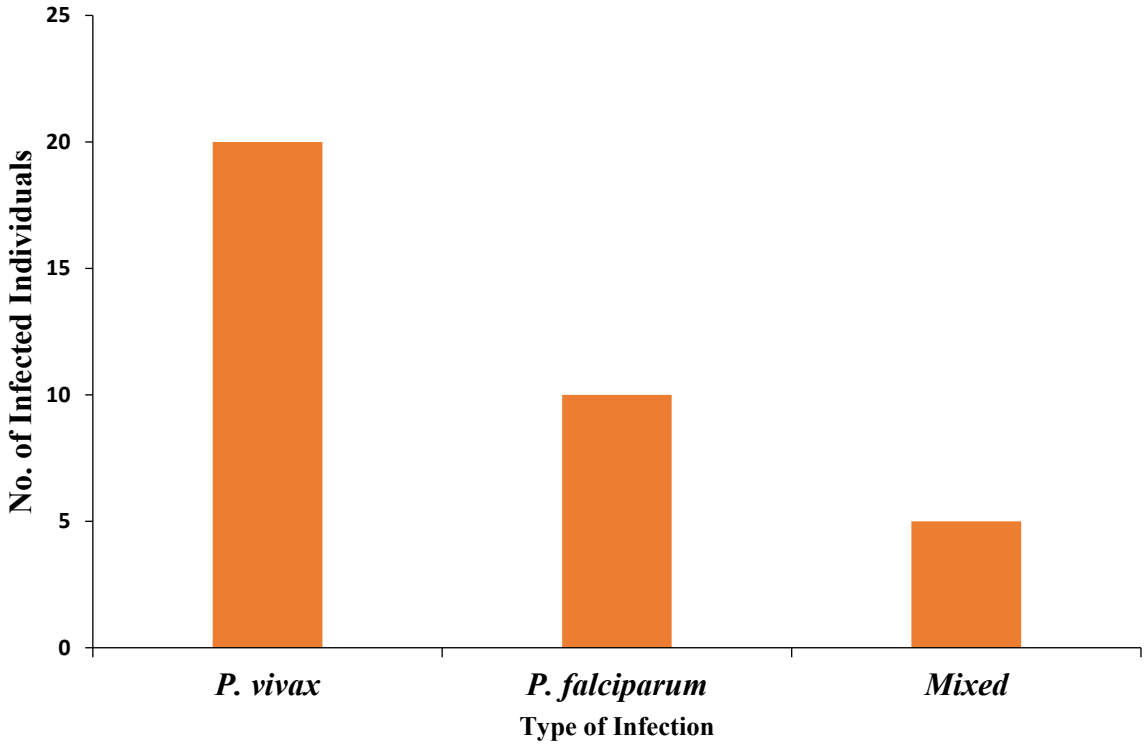

Fig. 3.1: Number of infected malarial cases according to type of infection.

3.2.2 Gender-wise malarial cases

Out of a total of 35 people who had malaria infection, 20 patients were male accounting for 52.147 % of cases, and 15 patients were females with 42.923 % of cases (Table 3.2 and Fig.3.2).

Table 3.2: Gender-wise incidence of malaria infection.

Gender	No. of Infected Individuals	Percentage
Male	20	57.1%
Female	15	42.9%

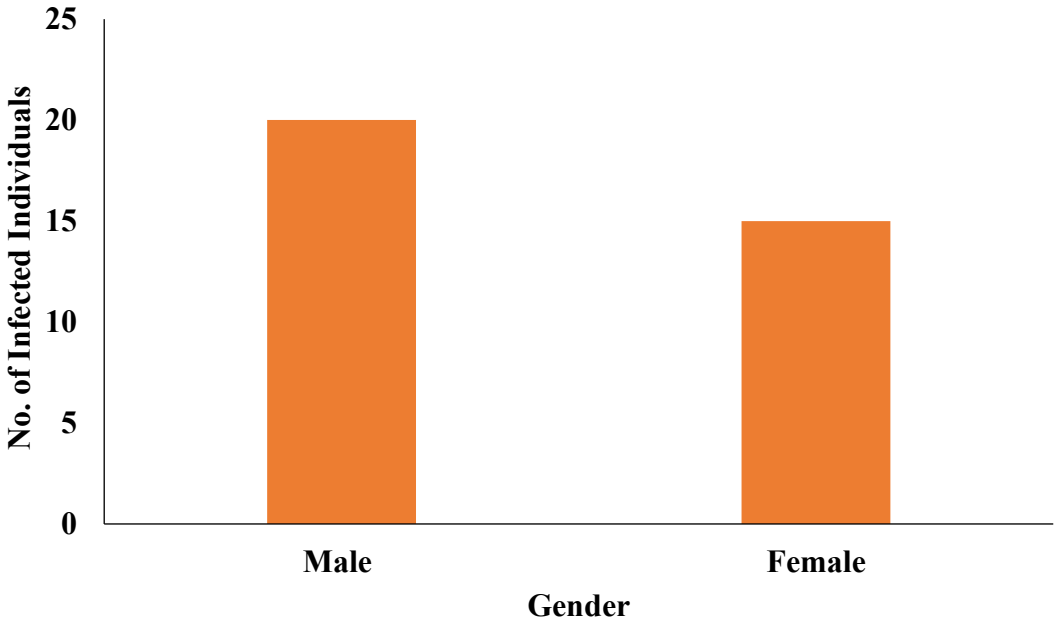

Fig.3.2. Number of infected malarial cases according to gender

3.3. DNA Isolation

DNA extraction was carried out for the malaria-positive blood sample, yielding an average DNA concentration of approximately 50ng/µl. Additionally, the isolated DNA showed a purity value of 1.8 across all samples. These findings highlight the successful extraction process and quality of DNA which provides a strong foundation for genetic analysis (Fig.3.3).

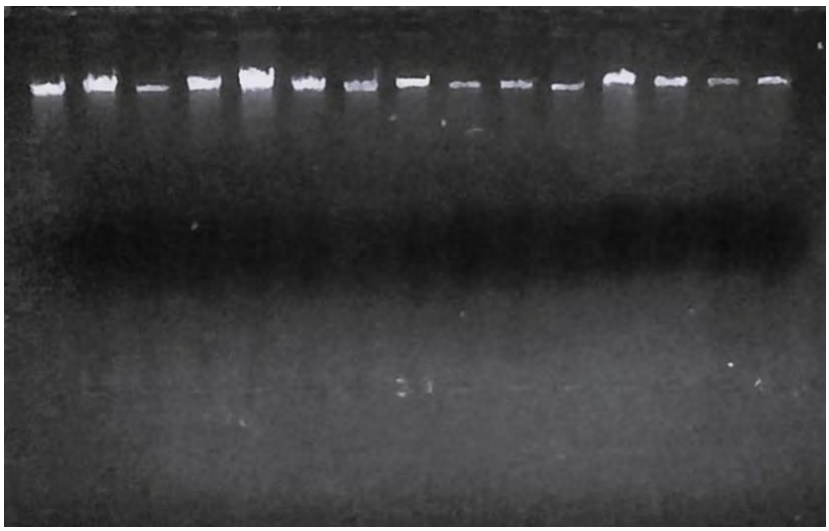

Fig.3.3. Agarose gel electrophoresis of extracted DNA samples.

3.4. Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) was performed to amplify the DNA samples of selected patients who were infected with *Plasmodium falciparum*. The PCR products for the Exon 5 of the GYPB gene was 373 bp. The conformation of the amplicon was done by running the 2% of agarose gel electrophoresis and was observed in the Gel Documentation system as shown in Fig. 3.4.

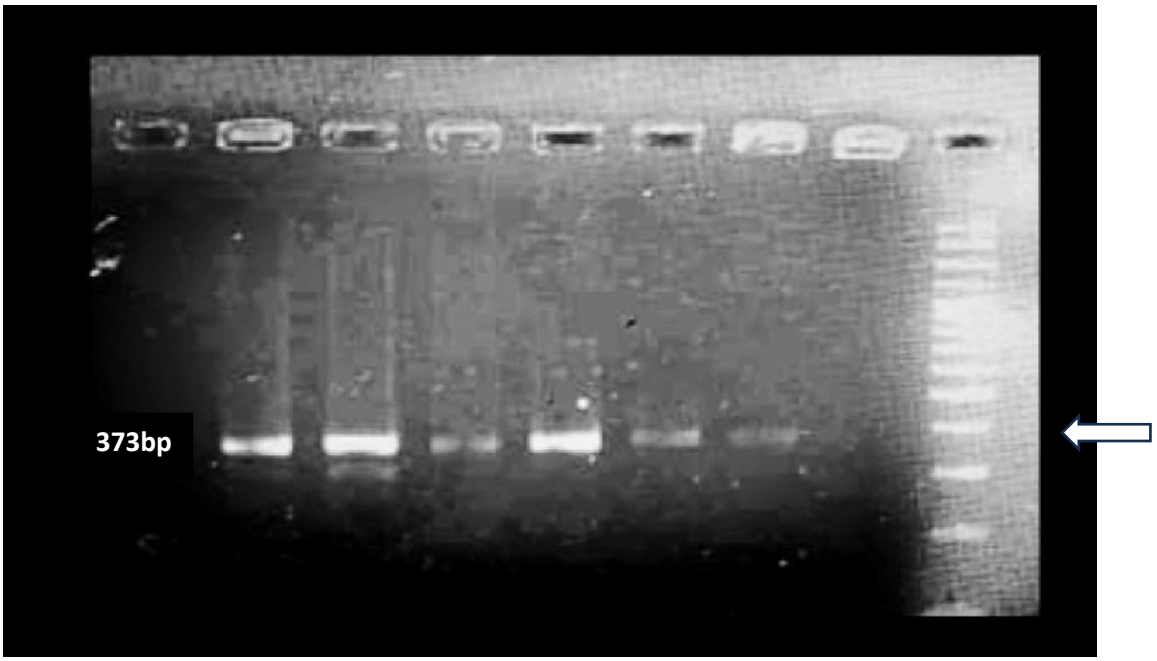

Fig.3.4: Gene amplification PCR products of GYPB exon 5.

3.5. Genetic Analysis

Polymorphism in the GYPB gene is known to affect RBCs Glycophorin-B specificity and binding to the ligand of pathogen consequently decreasing invasion efficiency and may cause susceptibility to malaria. Exon 5 of GYPB (Ensemble Transcript I.D: ENSE00002086685), which codes for the transmembrane domain of the GPB protein receptor, with a 373 bp length was selected to screen mutations and polymorphism. The GYPB gene having the Ensemble Transcript I.D: ENST00000502664.6 was used as the standard sequence for sequence alignment. The analyzed sample sequences are explained below.

3.6. Mutation Analysis

Table 3.3 represents the detailed results of mutation analysis of GYPB gene exon 5. Through the Sanger Sequencing of sample ME5.32, a reported heterozygous mutation (c.272C>A) was found in the exonic region of the GYPB gene's exon 5. The mutation results in an alteration of

the amino acid sequence in the protein structure. Furthermore, splice site changes occur, and protein features might be affected. The mutation was analyzed and evaluated by Polyphen software and was identified as a benign mutation. The sequencing chromatogram is shown in Fig 3.5.

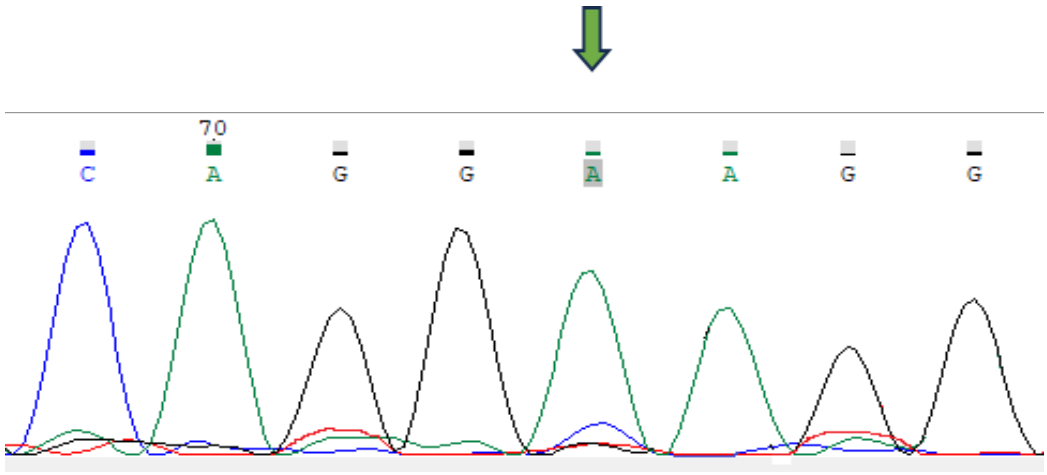

Fig.3.5 Sequencing chromatogram showing single nucleotide change at position c.270C>A in exon 5 of GYPB gene of Sample ME5.32.

Another heterozygous mutation, in sample ME5.32, which was not reported previously (cDNA.397A>G), was found in the 3'UTR region of the GYPB gene and resulted in changes in the splice site. The mutation was analyzed and evaluated by Polyphen software and was identified as a benign mutation. The sequencing chromatogram is shown in Fig 3.6.

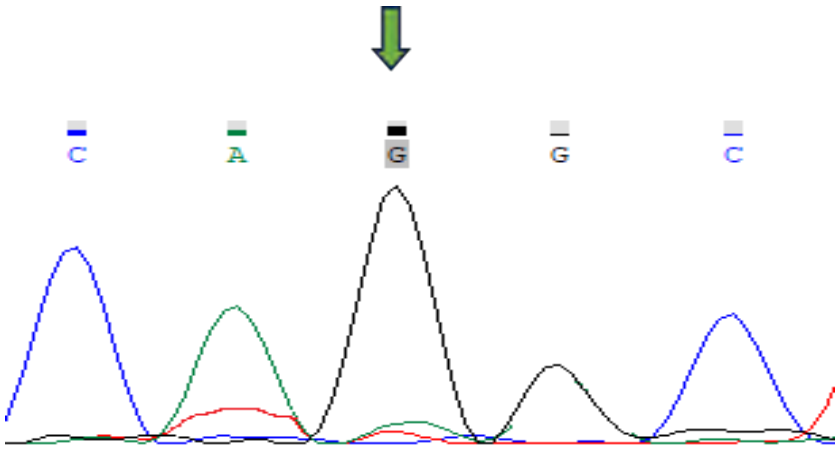

Fig.3.6 Sequencing chromatogram showing single nucleotide change at position c.DNA.397A>G in 3'UTR of GYPB gene of Sample ME5.32.

A heterozygous mutation, in sample ME5.24, that was not reported previously (cDNA.397A>G), was found in the 3'UTR region of the GYPB gene and resulted in changes in the splice site The sequencing chromatogram is shown in Fig 3.7.

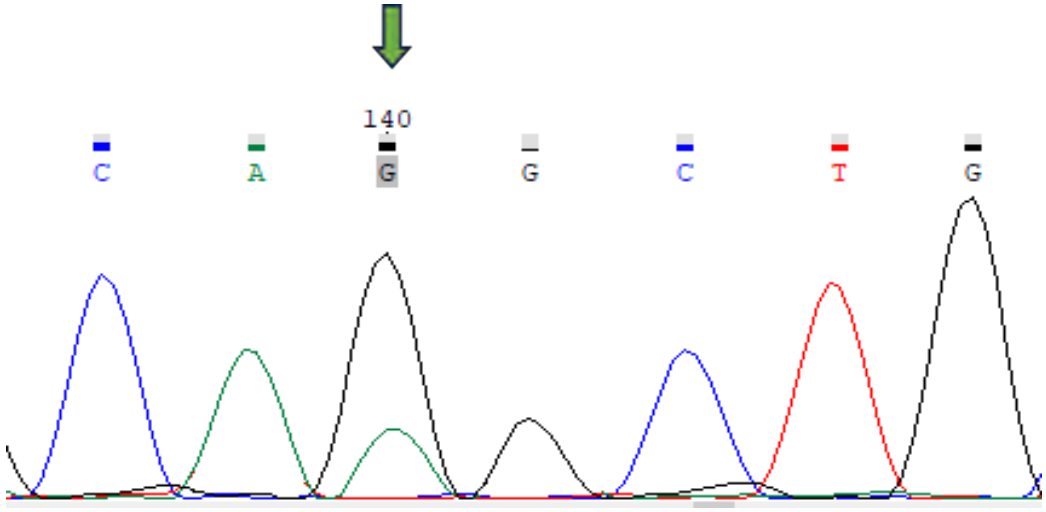

Fig.3.7 Sequencing chromatogram showing single nucleotide change at position c.DNA.397A>G in 3'UTR of GYPB gene of Sample ME5.24.

A heterozygous mutation, in sample ME5.21, that was not reported previously (cDNA.442C>T), was found in the 3'UTR region of the GYPB gene and resulted in changes in the splice site. The sequencing chromatogram is shown in Fig 3.8.

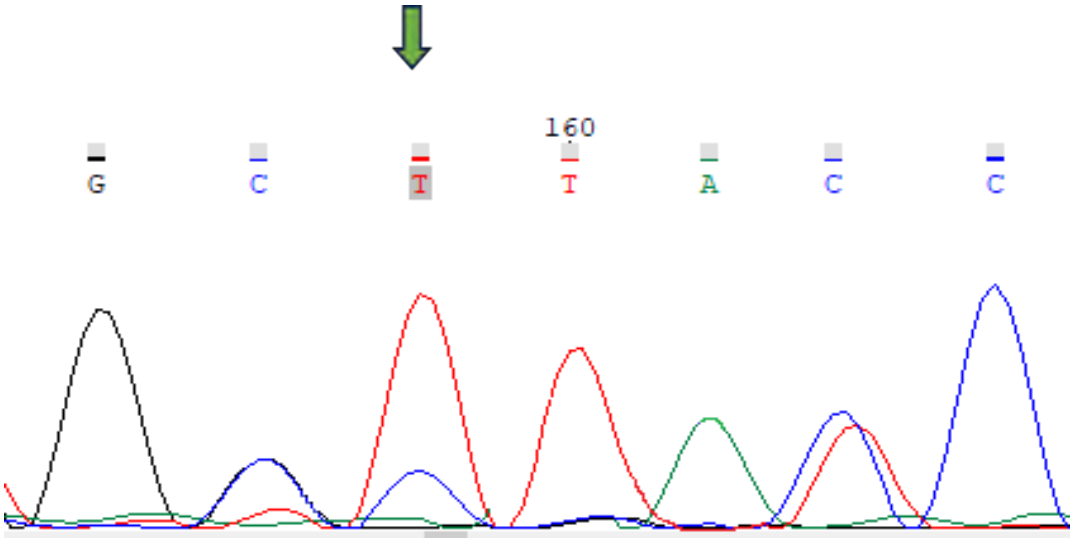

Fig.3.8 Sequencing chromatogram showing single nucleotide change at position c.DNA.422C>T in 3'UTR of GYPB gene of Sample ME5.21.

A reported mutation, in sample ME5.21, (g.41355T>C) was found in the intronic region of the GYPB gene. The mutation results in an alteration of the amino acid sequence in the protein structure. Furthermore, splice site changes occur, and protein features might be affected. The sequencing chromatogram is shown in Fig 3.9.

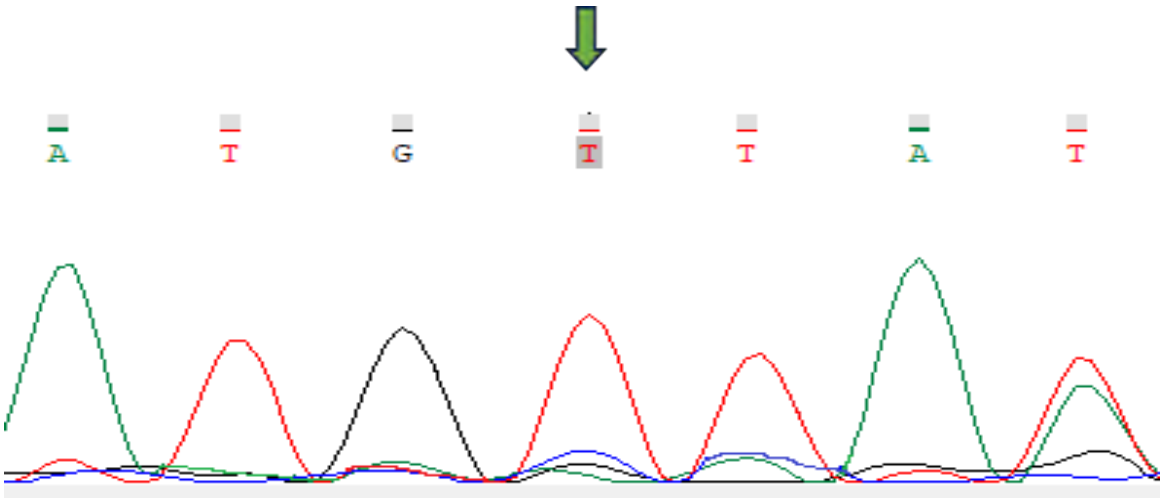

Fig.3.9 Sequencing chromatogram showing single nucleotide change at position g.41355T>C in the intronic region of GYPB gene of Sample ME5.21

A heterozygous mutation, in sample ME5.18, which was not reported previously (cDNA.443T>C), was found in the 3'UTR region of the GYPB gene and resulted in changes in the splice site. The sequencing chromatogram is shown in Fig 3.10.

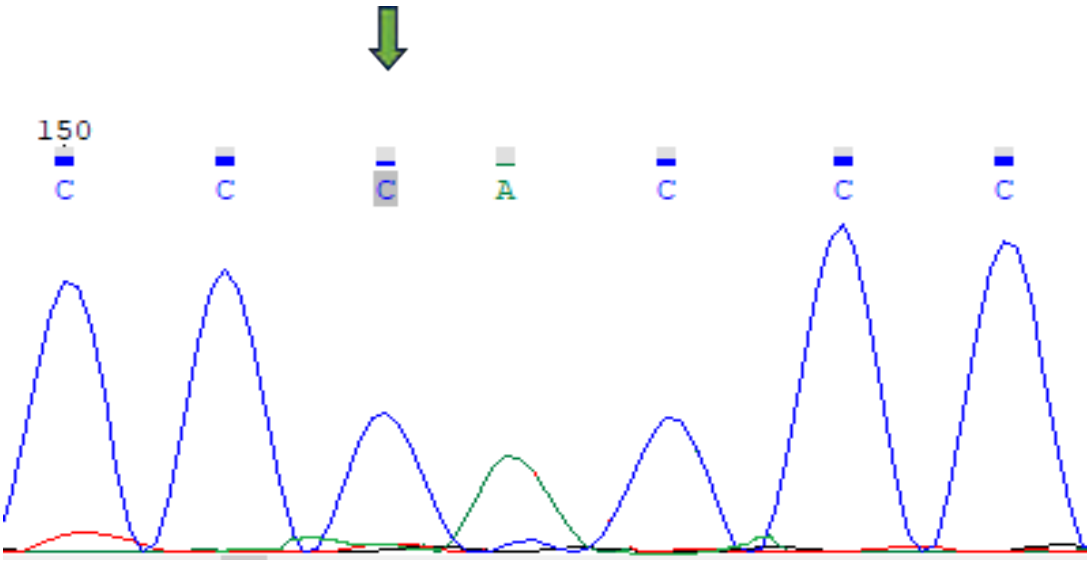

Fig.3.10 Sequencing chromatogram showing single nucleotide change at position c.DNA.433T>C in 3'UTR of GYPB gene of Sample ME5.18

Another heterozygous mutation, in sample ME5.18, which was not reported previously (cDNA.457_458insG), was found in the 3'UTR region of the GYPB gene and resulted in changes in the splice site and change in the signal of Poly A tail as well. The chromatogram is shown in Fig 3.11.

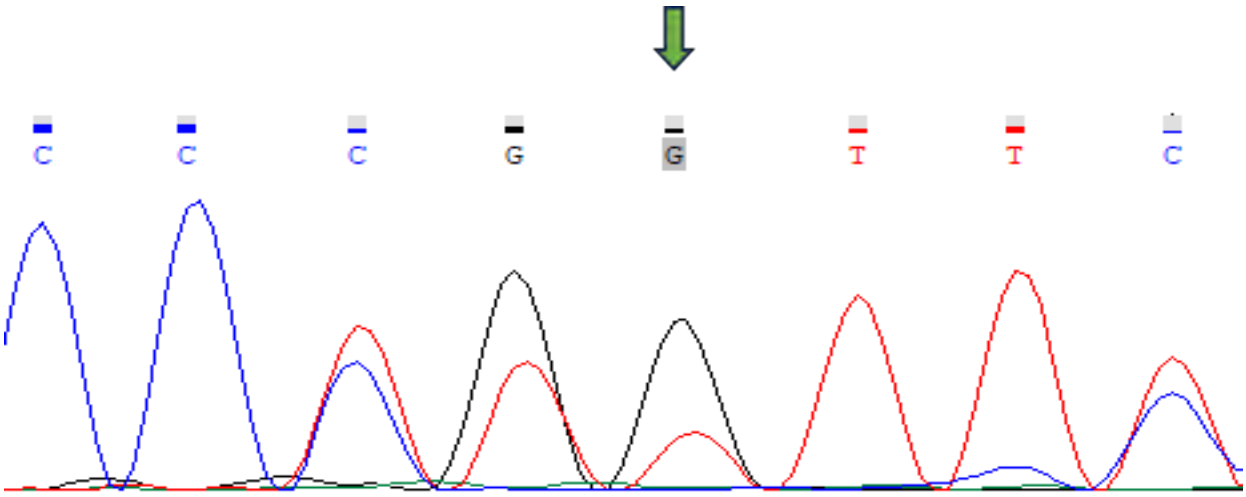

Fig.3.11 Sequencing chromatogram showing insertion at position cDNA.457_458insG in 3'UTR of GYPB gene of Sample ME5.18

Another reported mutation, in sample ME5.18, (g.41355T>C) was found in the intronic region of the GYPB gene. The mutation results in an alteration of the amino acid sequence in the protein structure. Furthermore, splice site changes occur, and protein features might be affected. The sequencing chromatogram is shown in Fig 3. 12.

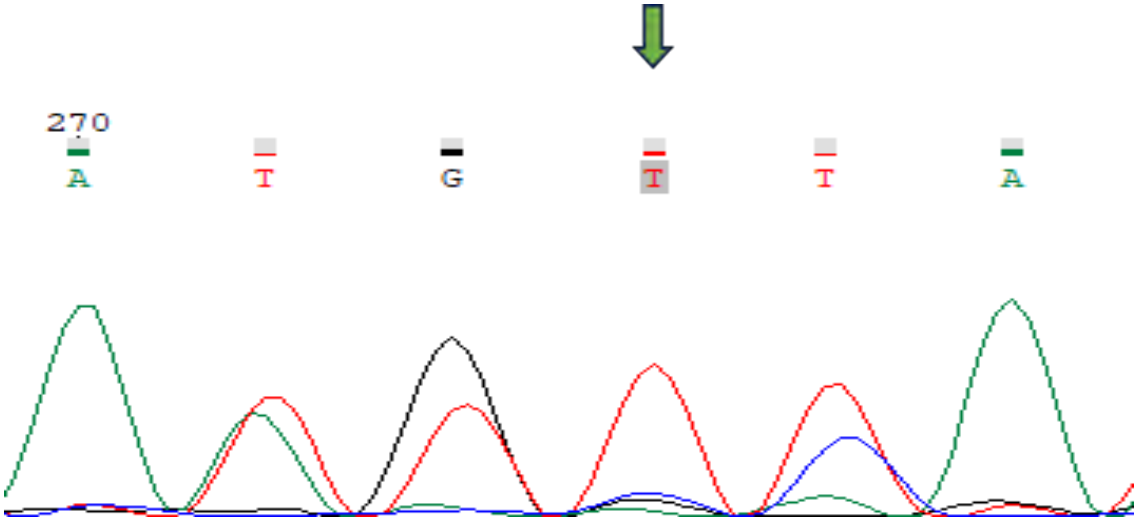

Fig.3.12 Sequencing chromatogram showing single nucleotide change at position g.41355T>C in the intronic region of GYPB gene of Sample ME5.18

Table 3.3 Identified Polymorphisms and Mutations in Malaria-Positive Patients

Sample ID	Genomic location	Physical Location	AA change	rs.ID	Mutation Tester	Status
ME5.32	(c.272C>A)	chr4:144917456	A91E	rs369309353	Polymorphism	Reported
ME5.32	g.144462A>G	chr4:144917383	N/A	N/A	Polymorphism	Novel
ME5.24	g.144462A>G	chr4:144917383	N/A	N/A	Polymorphism	Novel
ME5.21	g.144507C>T	chr4:144917338	N/A	N/A	Polymorphism	Novel
ME5.21	g.41355T>C	chr4:145020490	N/A	rs6537251	Polymorphism	Reported
ME5.18	g.144498T>C	chr4:144917347	N/A	N/A	Polymorphism	Novel
ME5.18	g.144522_144523insG	chr4:144917322_144917323	N/A	N/A	Polymorphism	Novel
ME5.18	g.41355T>C	chr4:145020490	N/A	rs6537251	Polymorphism	Reported

DISCUSSION

Malaria is a vector-linked ailment that continues to pose substantial health challenges to individuals residing in its endemic areas (Varo *et al* ., 2020). The infection develops when Anopheles mosquitoes transmit Plasmodium species through its bite (Meibalan *et al*., 2017). The WHO considers malaria a deadly and ubiquitous disease (WHO, 2020). Malaria is found in 86 tropical and subtropical nations, putting 50% of the world's population at risk. These cover Africa, the Mediterranean, the Western Pacific, Southeast Asia, and Pakistan (WHO, 2022).

Out of the 120 Plasmodium species, only 6 have been identified as the causative agents of malaria. Plasmodium falciparum is responsible for global malaria fatalities, while Plasmodium vivax often causes simple malaria but can nevertheless result in severe sickness in specific instances. (Naing *et al*., 2014). Erythrocyte invasion by P. falciparum depends on glycophorins. They are receptors for Erythrocytes Binding Antigens, the main parasite ligands. Glycophorins are proteins and Sialic acid (Orlandi *et al*., 1992). Although O and N linked Sialic acid can bind to the parasite ligand, their presence does not guarantee optimal binding. Thus, effective binding requires a certain amino acid sequence (Sim *et al*., 1994). Due to the necessity of this particular amino acid sequence, parasite-binding regions of red blood cell receptors will vary genetically. These changes affect binding strength and selectivity. Different Plasmodium falciparum ligands can affect red blood cell invasion. (Jiang *et al* ., 2011).

Multiple methods can identify genetic variants or polymorphisms, including enzymatic and direct methods like Differential Sequencing and RFLP (Kristensen *et al*., 2001). Direct sequencing of PCR products has advantages, but the biggest challenge is resolving sequencing errors caused by extraneous primers, nucleotides, or numerous products. (Omedo *et al* ., 2011).

This study applied direct sequencing of PCR results to identify allelic variants of GYPB's exon 5 throughout the Khushab population, which is known to be malaria endemic. Specific primers were designed using the Basic Local Alignment Search (BLAST) to confirm the amplification and sequencing of genes. The amplification of Exon 5 of the GYPB gene is crucial because it encodes the transmembrane domain of the protein, which plays a vital role in binding with the cell membrane of red blood cells.

The primary objective of this study was to identify genetic variations, such as insertions, deletions, and single nucleotide polymorphisms (SNPs), which are changes in the DNA

sequence. Single nucleotide polymorphisms (SNPs) constitute approximately 90% of the overall genetic variety in humans, making them the most abundant kind of genetic variation (Crawford *et al*., 2005). Multiple heterozygous sites, exhibiting Single Nucleotide Polymorphism (SNP), were detected in the exon 5 region of GYPB. The polymorphisms were found in both the exonic and intronic areas of the GYPB exon 5, with the intronic portions being more prevalent. Intronic regions lack protein-coding activity and so do not experience significant selection pressures to preserve their sequence integrity. This aligns with expectations (Llopart *et al*., 2002).

The sequencing approach employed in this study, along with the diploid human genome, allows for the simultaneous identification of both alleles of the GYPB gene. This phenomenon is clearly observable when there are two distinct peaks of varying colors that correspond to each of the two bases in instances of heterozygosity at specific sites in the glycophorin gene (Omedo *et al*., 2011).

Mutations at splice sites that direct intron exclusion during transcription could affect the gene product and result in a non-functional protein (Holland et al., 2001). Therefore, these changes must be carefully examined. The intronic and 3'UTR sections of samples ME5.32, ME5.24, ME5.21, and ME5.18 have 7 single nucleotide polymorphisms. A similar study was done in Kilifi, Kenya, where malaria is common. Three main glycophorin genes—GYPA, GYPB, and GYPC—were sequenced by direct PCR. Malaria-infected blood was used (Omedo *et al*., 2011). Their investigation found the intronic region to be the most polymorphic, supporting our findings.

Sample ME5.18 had a GYPB gene intronic region 4-5 insertion. Sample ME5.32's exonic region had a non-synonymous mutation. Synonymous replacements affect codons but not amino acids, while non-synonymous substitutions change amino acids. Silent mutations are synonymous substitutions (Wang *et al*., 2017). Non-synonymous mutations can change protein structure and cause function loss, gain, or stop codons. These changes dramatically affect protein function (Chu *et al*., 2019). Truncating the protein reduces its function, which harms the body (Omedo *et al*., 2011).

An analysis was conducted on the modification in sample ME5.32, and the findings indicate that this polymorphism leads to alterations in both the amino acid sequence and splicing site, potentially affecting the protein sequence. After conducting a more in-depth examination of the Polyphen2 software, it was determined that this specific SNP is classified

as a benign mutation. The P. falciparum parasite primarily utilises glycophorins as its receptors. Molecular mutations within the gene, especially in the regions responsible for interacting with ligands, can impact the effectiveness of binding and consequently the severity of the sickness (Omedo *et al*., 2011).

Conclusion and Recommendations

Through a comprehensive evaluation of genetic variants within the scope of this research, in comparison with the analogous findings in the existing literature, it becomes evident that identified variants of GYPB were not previously documented in the Pakistani population. All the exonic and intronic polymorphisms discovered in our study were not previously identified. The multiple polymorphisms and genetic variations in the GYPB gene indicate that the gene is rapidly evolving over the course of time. Furthermore, a comprehensive exploration of the genetic variations and polymorphism in the GYPB gene in the Pakistani population is essential to understand the role of this gene is malarial resistance or severity.

REFERENCES

Autino, B., Corbett, Y., Castelli, F., & Taramelli, D. (2012). Pathogenesis of malaria in tissues and blood. *Mediterranean journal of hematology and infectious diseases*, *4*(1).

Basu, S., & Sahi, P. K. (2017). Malaria: An Update. *Indian journal of pediatrics*, *84*(7), 521–528.

Caputo, A., & Garavelli, P. L. (2016). Climate, environment, and transmission of malaria. *Infez Med*, *2*, 93-104.

Chu, D., & Wei, L. (2019). Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation. BMC cancer, 19(1), 1-12.

Cowman, A. F., & Crabb, B. S. (2006). Invasion of red blood cells by malaria parasites. *Cell*, *124*(4), 755-766.

Cowman, A. F., Healer, J., Marapana, D., & Marsh, K. (2016). Malaria: biology and disease. Cell, 167(3), 610-624.

Crawford, D. C., & Nickerson, D. A. (2005). Definition and clinical importance of haplotypes. Annu. Rev. Med., 56, 303-320.

Daniels, G. (2013). Knops blood group system and the Cost antigens. *Human Blood Groups, 3rd edn (ed. by G. Daniels)*, 439-448.

Dean, L., & Dean, L. (2005). *Blood groups and red cell antigens* (Vol. 2). Bethesda: NCBI.

Dondorp, A. M., Pongponratn, E., & White, N. J. (2004). Reduced microcirculatory flow in severe falciparum malaria: pathophysiology and electron-microscopic pathology. *Acta tropica*, *89*(3), 309-317.

Doumbo, O. K., Thera, M. A., Koné, A. K., Raza, A., Tempest, L. J., Lyke, K. E., ... & Rowe, J. A. (2009). High levels of Plasmodium falciparum rosetting in all clinical forms of severe malaria in African children. *The American journal of tropical medicine and hygiene*, *81*(6), 987.

Garrido-Cardenas, J. A., González-Cerón, L., Manzano-Agugliaro, F., & Mesa-Valle, C. (2019). Plasmodium genomics: an approach for learning about and ending human malaria. *Parasitology Research*, *118*(1), 1–27.

Gilson, P. R., Chisholm, S. A., Crabb, B. S., & de Koning-Ward, T. F. (2017). Host cell remodeling in malaria parasites: a new pool of potential drug targets. *International journal for parasitology*, *47*(2-3), 119-127.

Grau, G. E. R., & Craig, A. G. (2012). Cerebral malaria pathogenesis: revisiting parasite and host contributions. *Future microbiology*, *7*(2), 291-302.

Harbach, R. E. (1994). Review of the internal classification of the genus Anopheles (Diptera: Culicidae): the foundation for comparative systematics and phylogenetic research. *Bulletin of Entomological Research*, *84*(3), 331-342.

Holland, J. B., Helland, S. J., Sharopova, N., & Rhyne, D. C. (2001). Polymorphism of PCR-based markers targeting exons, introns, promoter regions, and SSRs in maize and introns and repeat sequences in oat. Genome, 44(6), 1065-1076.

Huang, C. H., & Blumenfeld, O. O. (1995). MNSs blood groups and major glycophorins: molecular basis for allelic variation. In *Molecular basis of human blood group antigens* (pp. 153-188). Boston, MA: Springer US.

Jiang, H., Li, N., Gopalan, V., Zilversmit, M. M., Varma, S., Nagarajan, V., ... & Su, X. Z. (2011). High recombination rates and hotspots in a Plasmodium falciparum genetic cross. Genome biology, 12(4), 1-15.

Karim, A. M., Yasir, M., Ali, T., Malik, S. K., Ullah, I., Qureshi, N. A., ... & Jin, H. J. (2021). Prevalence of clinical malaria and household characteristics of patients in tribal districts of Pakistan. *PLoS Neglected Tropical Diseases*, *15*(5), e0009371.

Kristensen, V. N., Kelefiotis, D., Kristensen, T., & Børresen-Dale, A. L. (2001). High-throughput methods for detection of genetic variation. Biotechniques, 30(2), 318-332.

Li, X., Marinkovic, M., Russo, C., McKnight, C. J., Coetzer, T. L., & Chishti, A. H. (2012). Identification of a specific region of Plasmodium falciparum EBL-1 that binds to host receptor glycophorin B and inhibits merozoite invasion in human red blood cells. *Molecular and biochemical parasitology*, *183*(1), 23-31.

Llopart, A., Comeron, J. M., Brunet, F. G., Lachaise, D., & Long, M. (2002). Intron presence–absence polymorphism in Drosophila driven by positive Darwinian selection. Proceedings of the National Academy of Sciences, 99(12), 8121-8126.

Looareesuwan, S., & Wilairatana, P. (1994). Clinical tropical medicine I.

Lyons, E. A., Scheible, M. K., Sturk-Andreaggi, K., Irwin, J. A., & Just, R. S. (2013). A high-throughput Sanger strategy for human mitochondrial genome sequencing. BMC genomics, 14, 1-16.

Mawson, A. R. (2013). The pathogenesis of malaria: a new perspective. *Pathogens and global health*, *107*(3), 122-129.

McGhee, R. B. (1953). The infection by Plasmodium lophurae of duck erythrocytes in the chicken embryo. *The Journal of Experimental Medicine*, *97*(6), 773-782.

McMorrow, M. L., Masanja, M. I., Abdulla, S. M., Kahigwa, E., & Kachur, S. P. (2008). Challenges in routine implementation and quality control of rapid diagnostic tests for malaria–Rufiji District, Tanzania. *The American journal of tropical medicine and hygiene*, *79*(3), 385.

Mei, Y., Liu, Y., & Ji, P. (2021). Understanding terminal erythropoiesis: An update on chromatin condensation, enucleation, and reticulocyte maturation. Blood Reviews, 46, 100740.

Meibalan, E., & Marti, M. (2017). Biology of malaria transmission. *Cold Spring Harbor Perspectives in Medicine*, *7*(3), a025452.

Milner, D. A. (2018). Malaria pathogenesis. *Cold Spring Harbor perspectives in medicine*, *8*(1), a025569.

Molina-Franky, J., Patarroyo, M. E., Kalkum, M., & Patarroyo, M. A. (2022). The cellular and molecular interaction between erythrocytes and Plasmodium falciparum merozoites. Frontiers in cellular and infection microbiology, 12, 816574. Mouchet, J. (2004). *Biodiversité du paludisme dans le monde*. John Libbey Eurotext.

Mukry, S. N., Saud, M., Sufaida, G., Shaikh, K., Naz, A., & Shamsi, T. S. (2017). Laboratory diagnosis of malaria: comparison of manual and automated diagnostic tests. *Canadian Journal of Infectious Diseases and Medical Microbiology*, *2017*.

Mukry, S. N., Saud, M., Sufaida, G., Shaikh, K., Naz, A., & Shamsi, T. S. (2017). Laboratory diagnosis of malaria: comparison of manual and automated diagnostic tests. Canadian Journal of Infectious Diseases and Medical Microbiology, 2017.

Naing, C., Whittaker, M. A., Nyunt Wai, V., & Mak, J. W. (2014). Is Plasmodium vivax malaria a severe malaria?: a systematic review and meta-analysis. PLoS neglected tropical diseases, 8(8), e3071.

Omedo, I. A. (2013). Characterization Of Erythrocyte Receptor Polymorphisms In A Malaria Endemic Population In Kilifi, Kenya (Doctoral dissertation, University of Nairobi).

Orlandi, P. A., Klotz, F. W., & Haynes, J. D. (1992). A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of Plasmodium falciparum recognizes the terminal Neu5Ac (alpha 2-3) Gal-sequences of glycophorin A. The Journal of cell biology, 116(4), 901-909.

Pasvol, G., Wainscoat, J. S., & Weatherall, D. J. (1982). Erythrocytes deficient in glycophorin resist invasion by the malarial parasite Plasmodium falciparum. Nature, 297(5861), 64-66.

Peterson, D. S., & Wellems, T. E. (2000). EBL-1, a putative erythrocyte binding protein of Plasmodium falciparum, maps within a favored linkage group in two genetic crosses. *Molecular and biochemical parasitology*, *105*(1), 105-113.

Peterson, D. S., Miller, L. H., & Wellems, T. E. (1995). Isolation of multiple sequences from the Plasmodium falciparum genome that encodes conserved domains homologous to those in erythrocyte-binding proteins. *Proceedings of the National Academy of Sciences*, *92*(15), 7100-7104.

Price, R. N., Douglas, N. M., & Anstey, N. M. (2009). New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance. *Current opinion in infectious diseases*, *22*(5), 430-435.

Reid, M. E. (1994). Some concepts relating to the molecular genetic basis of certain MNS blood group antigens. Transfusion Medicine, 4(2), 99-111.

Rowe, J. A., Claessens, A., Corrigan, R. A., & Arman, M. (2009). Adhesion of Plasmodium falciparum-infected erythrocytes to human cells: molecular mechanisms and therapeutic implications. *Expert reviews in molecular medicine*, *11*, e16.

Salinas, N. D., Paing, M. M., & Tolia, N. H. (2014). Critical glycosylated residues in exon three of erythrocyte glycophorin A engage Plasmodium falciparum EBA-175 and define receptor specificity. *MBio*, *5*(5), 10-1128.

Sim, B. K. L., Chitnis, C. E., Wasniowska, K., Hadley, T. J., & Miller, L. H. (1994). Receptor and ligand domains for invasion of erythrocytes by Plasmodium falciparum. Science, 264(5167), 1941-1944.

Stevenson, M. M., & Riley, E. M. (2004). Innate immunity to malaria. *Nature Reviews Immunology*, *4*(3), 169-180.

Talapko, J., Škrlec, I., Alebić, T., Jukić, M., & Včev, A. (2019). Malaria: the past and the present. *Microorganisms*, *7*(6), 179.

Talapko, J., Škrlec, I., Alebić, T., Jukić, M., & Včev, A. (2019). Malaria: the past and the present. Microorganisms, 7(6), 179.

Tangpukdee, N., Duangdee, C., Wilairatana, P., & Krudsood, S. (2009). Malaria diagnosis: a brief review. *The Korean journal of parasitology*, *47*(2), 93.

Taramelli, D., Recalcati, S., Basilico, N., Olliaro, P., & Cairo, G. (2000). Macrophage preconditioning with synthetic malaria pigment reduces cytokine production via heme iron-dependent oxidative stress. *Laboratory Investigation*, *80*(12), 1781-1788.

Tarazona-Santos, E., Castilho, L., Amaral, D. R., Costa, D. C., Furlani, N. G., Zuccherato, L. W., ... & Lustigman, S. (2011). Population genetics of GYPB and association study between GYPB* S/s polymorphism and susceptibility to P. falciparum infection in the Brazilian Amazon. *PloS one*, *6*(1), e16123.

Treatment of Malaria: Guidelines for Clinicians (United States), CDC (2023)

Tuteja, R. (2007). Malaria is a global disease. FEBS J, 274(18), 4669.

Urban, B. C., Ing, R., & Stevenson, M. M. (2005). Early interactions between blood-stage plasmodium parasites and the immune system. *Immunology and Immunopathogenesis of Malaria*, 25-70.

Varo, R., Chaccour, C., & Bassat, Q. (2020). Update on malaria. Medicina Clínica (English Edition), 155(9), 395-402.

Wang, H. Y., Tang, H., Shen, C. K. J., & Wu, C. I. (2003). Rapidly evolving genes in human. I. The glycophorins and their possible role in evading malaria parasites. *Molecular biology and evolution*, *20*(11), 1795-1804.

Wang, S., Li, L., Tao, R., & Gao, Y. (2017). Ion channelopathies associated genetic variants as the culprit for sudden unexplained death. Forensic science international, 275, 128-137.

Wassmer, S. C., & Carlton, J. M. (2016). Glycophorins, Blood Groups, and Protection from Severe Malaria. *Trends in parasitology*, *32*(1), 5–7. https://doi.org/10.1016/j.pt.2015.11.006

White, N. J., Pukrittayakamee, S., Hien, T. T., Faiz, M. A., Mokuolu, O. A., & Dondorp, A. M. (2014). Malaria. *The Lancet*, *383*(9918), 723-735.

WHO Study Group on the Implementation of the Global Plan of Action for Malaria Control & World Health Organization. (1993). Implementation of the global malaria control strategy: report of a WHO Study Group on the Implementation of the Global Plan of Action for Malaria Control 1993-2000 [meeting held in Geneva from 8 to 12 February 1993]. World Health Organization.

Wilson, M. L. (2013). Laboratory diagnosis of malaria: conventional and rapid diagnostic methods. *Archives of Pathology and Laboratory Medicine*, *137*(6), 805-811.

World Health Organization, 2010. Guidelines for the treatment of malaria. Geneva: WHO

World Health Organization. (2022). World malaria report 2022. World Health Organization.

World Health Organization. World Malaria Report 2020:20 Years of Global Progress and Challenges. Geneva: World Health Organization (2020).

Zekar, L., & Sharman, T. (2020). Plasmodium falciparum malaria

ANNEXURE I

Table 1 Chemical concentration and composition used in DNA extraction.

S.No	Chemical	Concentration
1	Sol A	500 µL
2	Solution B	400 µL
3	Sodium Dodecyl Sulphate (SDS)	25 µL
4	Proteinase Kinase (PK)	5 µL
5	Solution C+D	500 µL
6	Solution D	500 µL
7	Isopropyl Alcohol	500 µL
8	Sodium Acetyl	60 µL
9	T.E Buffer	200 µL

Table 2 Composition of Agarose gel and other required chemicals

S.No	Solutions	Composition
1	1% Agarose Gel (50 mL)	1X TBE Buffer (50 mL)
		Agarose (0.5 g)
		Ethidium Bromide (2 µL)
2	2% Agarose Gel (50 mL)	10X TBE Buffer (5 mL)
		Agarose (1.0 g)
		Ethidium Bromide (5 µL)
		Distilled water (45 mL)
3	Gel Preparation Buffer (10X TBE)	Boric Acid (27.5 g)
		EDTA (3.6 g)
		Tris (54 g)
		Deionized water (500 mL)
4	Gel Running Buffer (1X TBE)	10X TBE Buffer (1 part)
		Distilled water (9 parts)
5	Ethidium Bromide Solution (50 mL)	Autoclaved filter water (50 mL) Ethidium Bromide (0.5 g)

6	Loading Dye Solution (25 mL)	Autoclaved filter water (25 mL)
		Bromophenol blue (0.087 g)
		Sucrose (10 g)

Table 3 Chemical concentration and volume used in the reaction mixture

Chemicals	Concentration	Volume
Taq buffer	10X	2.5 µL
dNTPs	2.5 mM	2.5 µL
MgCl2	2.5 mM	2 µL
Forward Primer	10 pmol/µL	0.5 µL
Reverse Primer	10 pmol/µL	0.5 µL
DNA	>100 ng/µL	2 µL
Taq Polymerase	5 U/µL	0.5 µL
PCR Water		14.5 µL

ORIGINALITY REPORT

17%	10%	8%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.unn.edu.ng Internet Source	1%
2	Submitted to Nelson Mandela Metropolitan University Student Paper	1%
3	Submitted to MAHSA University Student Paper	1%
4	Submitted to University of Northumbria at Newcastle Student Paper	1%
5	Submitted to CSU, Hayward Student Paper	1%
6	Submitted to Chester College of Higher Education Student Paper	1%
7	journals.plos.org Internet Source	1%
8	Alan F. Cowman, Julie Healer, Danushka Marapana, Kevin Marsh. "Malaria: Biology and Disease", Cell, 2016	1%