UNIVERSITY OF GHANA

COLLEGE OF BASIC AND APPLIED SCIENCES

SCHOOL OF BIOLOGICAL SCIENCES

GENETIC VARIATION AND INVASION PHENOTYPES OF *PLASMODIUM FALCIPARUM* IN THE GAMBIA

BY

NGANYEWO NORA NGHOCHUZIE

(10857052)

A DISSERTATION SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES, UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF A DOCTOR OF PHILOSOPHY DEGREE IN MOLECULAR CELL BIOLOGY OF INFECTIOUS DISEASES

DEPARTMENT OF BIOCHEMISTRY, CELL, AND MOLECULAR BIOLOGY

MARCH 2024

DECLARATION

All experiments conducted and presented in this thesis were done by me, Nganyewo Nora Nghochuzie at the Medical Research Council Unit The Gambia at London School of Hygiene and Tropical Medicine and the Department of Biochemistry, Cell and Molecular Biology, University of Ghana under the supervision of Prof. Alfred Amambua-Ngwa, Dr. Lucas Amenga-Etego, and Prof. Gordon A. Awandare.

Nganyewo Nora Nghochuzie Date: 13 September 2024

(Candidate)

Prof. Alfred Amambua-Ngwa Date: 16 September 2024

(Supervisor)

Dr. Lucas Amenga-Etego Date: 16 September 2024

(Supervisor)

Prof. Gordon A. Awandare Date: 16 September 2024

(Supervisor)

ABSTRACT

Plasmodium falciparum is the deadliest of all *Plasmodium* species causing malaria in humans. The interactions between *P. falciparum* merozoite ligands and erythrocyte receptors during erythrocyte invasion is a vital stage in the parasite's life cycle, linked with the initiation of the clinical signs/symptoms of the disease. The availability of multiple alternative invasion pathways gives the parasite the luxury of switching between invasion pathways as an immune evasion mechanism and a means of escaping from control interventions probably due to polymorphisms in both ligand and receptor genes or changes in gene expression profiles. Additionally, the malaria disease has been suggested to have imposed evolutionary selection pressure on the human genome resulting in evolutionary changes in genes encoding both human receptors and parasite ligands. However, research on variations in parasite ligands and their corresponding receptors has received far less attention. Thus, the main aim of this study was to explore host-parasite genetic and transcriptomic correlates of *P. falciparum* erythrocyte invasion phenotypes, and malaria outcomes in The Gambia. Specifically, the genetic variations and transcriptional changes of *P. falciparum* invasion ligands in isolates with different invasion phenotypes were determined, sequence variations in parasite ligand genes and their corresponding erythrocyte receptors in the same population were also studied and lastly, the distribution of two key human glycophorin B receptor structural variants (denoted as Deletion 1 and 2) among malaria-infected individuals in The Gambia was described.

Following informed consent documentation, 2 ml of whole blood samples were collected from all age groups (one to seventy years) with confirmed malaria (severe and mild) in four health facilities across three regions in The Gambia, notably Basse (Upper River Region, mesoendemic), Brikama and Fajikunda (Western region, hypoendemic), and Edward Frances Small Teaching Hospital (Greater Banjul region, hypoendemic). Leukocyte-depleted infected red blood cells (iRBCs) were used for erythrocyte invasion experiments. The iRBCs were also

cultured to schizonts for RNA extraction and expression profiling of six invasion ligand genes: erythrocyte binding antigen (EBA) 175, EBA-181, reticulocyte homologue (Rh) 2b, Rh4, Rh5, and cytoadherence antigen 2 (Clag2), which are important players in the invasion machinery. To determine sequence variations in ligand and receptor genes, malaria-positive samples were used for targeted DNA sequencing of a broad range of 12 *P. falciparum* genes: EBA-175, EBA-181, EBA140, Clag2, Clag8, Rh4, Rh5, merozoite surface protein (MSP)1, MSP6, Duffy binding-like MSP (DBLMSP), erythrocyte binding-ligand 1 (EBL1), and surface-associated interspersed protein 4.2 (SURFIN4.2), and four human red cell receptors: glycophorin (GP) A, GPB, GPC, and complement receptor 1 (CR1), using the Oxford Nanopore GridIon platform. GPB deletion 1 and 2 were also genotyped from malaria-positive individuals using a multiplex PCR and restriction fragment length polymorphism-PCR approaches. Any sample with any or both of the deletions were further genotyped for sickle cell. R statistic packages were used for most analysis and the Shapiro-Wilk normality test was performed to determine if the distribution of the data was Gaussian, and variables that passed the test were subjected to parametric analysis; otherwise, non-parametric techniques were employed. Linkage disequilibrium was determined using the PLINK package and the Tajima's D test was used to estimate departures from neutral evolution in the selected parsite and host genes. The Hardy-Weinberg equation was used to determine the allele frequencies of GPB deletions while association analysis between the GPB deletions and severe malaria was undertaken using the Fisher's exact test.

A total of 90 clinical isolates were successfully cultured and phenotyped for the first part of the study. It was observed that *P. falciparum* clinical isolates in The Gambia predominantly use sialic-acid-independent pathways to invade the erythrocyte and the use of these pathways have increased from the year 2015 to 2022. Specifically, the data shows the increased expression of reticulocyte-binding homologue (Rh) protein family of parasite genes involved in sialic-acid-

independent pathways such as Rh5. Indeed, significantly positive correlations were observed between the expression of Rh2b, Rh4, and Rh5 known to be involved in the sialic-acid independent pathways. The *P. falciparum* isolates genotyped by nanopore sequencing showed moderate to high levels of within-host complexity of infection across sites. Interestingly, of the 12 genes sequenced, high inter-SNP linkage disequilibrium (LD) was observed only in the parasite DBLMSP and SURFIN 4.2 genes. LD was also strong for single nucleotide polymorphisms (SNPs) in the human GPB, GPC, and CR1 genes but not in GPA. Variants from these human loci (especially GPC) clustered individuals into two distinct sub-populations mostly shaped by two significant SNPs not previously described in The Gambia population. Moreover, Tajima's D analysis identified the parasite DBLMSP and SURFIN 4.2, and human host CR1, GPC, and GPB genes to be under balancing selection. A case (severe malaria) – control (mild malaria) univariable logistic regression models analysis identified a SNP in SURFIN 4.2 and some SNPs in the CR1 gene to be significantly associated with malaria susceptibility.

Additionaly, GPB deletions PCR genotyping showed high allele frequency of GPB deletion 1 (9.9 %) among malaria-positive individuals compared to deletion 2 (1.7 %). The Fulani and the Serehule ethnic groups had the highest frequency of GPB deletion 1 (14.8 %) and deletion 2 (2.5 %) respectively. Both deletions were more prevalent in Basse, where the Fulani ethnic group is dominant. However, both deletions were not associated with severe malaria propably due to small sample size of severe cases. Sickle cell genotyping among GPB deletion positive individulas suggested that the sickle cell and the GPB deletions may be co-evolving in population. Overall, this study has provided valuable insights into the parasite biology and the complex parasite's life cycle, while also identifying SNPs in both the parasite and host that could serve as potential targets for drugs and vaccines development.

DEDICATION

- To my baby whom I lost while writing this thesis. We were ready to welcome you, but I guess it wasn't the right time.
- In memory of my father Mr. Peyechu Clement Nganyewo, for being an amazing dad.
- To the entire Nganyewo family. I am proud to be part of you.
- To my husband Mr. Acho Fontebo Clovis for his unlimited support throughout this journey.

ACKNOWLEDGEMENT

My Ph.D. journey has been the most interesting, enjoyable, and confusing period of my life so far. I lack words to express my appreciation for all the amazing people who supported me and paved the way throughout this tough journey. For now, let me just say thank you for all your support throughout this journey.

Firstly, I would like to thank my supervisory team (Prof. Alfred Amambua-Ngwa, Dr. Lucas Amenga-Etego, and Prof. Gordon A. Awandare) uniquely for their support, great enthusiasm, freedom to explore, guiding me in designing and performing all experiments and criticism that were indispensable for the realization of this work. I appreciate your tolerance, your unwavering acceptance of me as a doctoral student, and your faith in me. You were not only supervisors but mentors, parents, and an entire package to push me through my career. I have learned a lot from you all and through you, my life has changed for the better.

To Dr. Eniyou C. Oriero, you have been like a big sister, a mentor, a friend, and a supervisor to me. Without you, I am sure I would have given up at some point. You listened to me at any time, opened your office at every little knock of mine, responded to every silly question of mine, advised me at my most confused moments, and encouraged me to push harder and keep on. I owe you a depth of gratitude and I can't thank you enough, but your reward is Divine.

I am grateful to every member of the malaria population biology (MPB) and genomics teams at the Medical Research Council Unit The Gambia at London School of Hygiene and Tropical Medicine and members of the Genomics and Bioinformatics group at WACCBIP for the training given to me and their relentless support. Special thanks goes to Dr. Abdul Karim Sesay, Mrs. Jarra Manneh, Dr Ajibola Olumide, Dr Majidah Hamid-Adiamoh, Dr Mary Oboh, Dr. Dominic S.Y. Amuzu, and Dr. Collins Misita Morang'a for grooming me through molecular research, Mr Simon Correa for introducing me to parasite culture, Fatoumatta Bojang and Dr

Ngoh Ines for introducing me to invasion experiments, Prof. Nuredin Mohammed, Mouhamadou Fadel, and Martha A. Demba for supporting with data analysis.

I am short of words to express my gratitude to my brother Mr Nganyewo Kenneth. No amount of appreciation can ever be enough for what you have done and are still doing for me. I think your purpose in this world was to take care of me because without you, there might have been no school for me and hence, no Ph.D.

To my mother, brothers, sisters, and Nganyewo Evodia, I am grateful for your encouragement, emotional, and spiritual support throughout this journey. Sincere gratitude goes to my husband Mr Acho Fontebo Clovis for not only encouraging me to apply for this Ph.D. fellowship but also accepting to live alone, experiencing all my ups, tolerating all my downs, and sharing all the hard times I went through during this journey.

I am grateful to all the participants and field workers who supported with sample collection, thereby making the successful completion of all experiments possible.

This research was conducted as a component of the Pan African Malaria Genetic Epidemiology Network (PAMGEN), an H3Africa project. It was also supported by the L'OREAL-UNESCO For Women in Science fellowship. Words fail me to appreciate this fully funded Ph.D. fellowship from PAMGEN especially Prof. Alfred A-Ngwa and team for granting me this opportunity through PAMGEN.

Last but not least, I am grateful to my mentor/father Dr. Nji Theobald. Papa as I always call you, I have not seen your type. It's been 7 years already and we are still counting. I don't even know where to start because tears wouldn't allow me to write. Thank you very much papa, your little daughter back then is making progress all because of you.

I will end by thanking Dr. Jonas Kengne for his care, support, confidence, trust, advice, and encouragement from my scholarship application till the completion of this Ph.D. journey.

THESIS OUTLINE

This thesis consists of 6 chapters, written in manuscript format whereby, the first chapter is the general introduction, hypotheses, objectives, and rationale for my work. The second chapter is the literature review while chapters 3 to 5 cover the results generated from the different objectives, Presented in manuscript format. Chapter 6 is the general discussion, conclusion, and recommendations. One paper from this thesis (chapter 3) has been published and the reference is listed below, while others are in preparation.

Nganyewo. N. N., Bojang. F., Oriero. E. C., Drammeh. F. N., Olumide. A., Mbye. H., Aminata J. S., Corea. S., Awandare. G. A., D'Alessandro. U., Amenga-Etego. L. N., and Amambua-Ngwa. A. (2023). Recent increase in low complexity polygenomic infections and sialic acid-independent invasion pathways in *Plasmodium falciparum* from Western Gambia. *Parasites and Vectors.* 16:309 https://doi.org/10.1186/s13071-023-05929-4

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µg	Microgram
µl	Microlitre
µM	Micromolar
ACT	Artemisinin Combination Therapy
AMA-1	Apical Membrane Antigen-1
ARDS	Acute Respiratory Distress Syndrome
BCF	Binary Variant Call Format
CM	Cerebral Malaria
COVID-19	Coronavirus Disease of 2029
CR1	Compliment Receptor-1
cRPMI	Complete Roswell Park Memorial Institute Media
CyRPA	Cysteine-rich Protector Antigen
DBS	Dry Blood Spots
DEL 1	Deletion 1
DEL 2	Deletion 2
DNA	Deoxyribonucleic acid
DUP4	Duplication 4
EBA	Erythrocyte Binding-like Antigen
EBL-1	Erythrocyte Binding Ligand
EDTA	Ethylenediaminetetraacetic Acid
EPCR	Endothelial Protein C Peceptor
G6PD	Glucose-6-Phosphate Dehydrogenase
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GP	Glycophorin

GTS	Global Technical Strategy
HbC	Haemoglobin C
HbE	Haemoglobin E
HbS	Haemoglobin S
HCT	Haematocrit
H*e*	Heterozygosity
ICAM1	Intercellular Adhesion Molecule 1
iRPMI	Incomplete Roswell Park Memorial Institute media
iTOL	Interactive Tree of Life
KDa	Kilodalton
LD	Linkage Disequilibrium
LLINs	Long-lasting Insecticidal Nets
MABL	Adhesive Erythrocyte Binding Protein
MOI	Multiplicity of Infection
MSP	Merozoite Surface Protein
MTRAP	Merozoite-Specific Thrombospondin Related Anonymous Protein
ONT	Oxford Nanopore Technology
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
PECAM	Platelet/endothelial Cell Adhesion Molecule
PfEMP-1	*Plasmodium falciparum* Erythrocyte Membrane Protein-1
PV	Parasitophorous Vacuole
RBCs	Red Blood Cells
Rh	Reticulocyte Homologue
RIFIN	Repetitive Interspersed Family

RPMI	Roswell Park Memorial Institute
RT-qPCR	Reverse Transcriptase Quantitative Polymerase Chain Reaction
SA	Sialic Acid
SAM	Sequence Alignment Map
SMA	Severe Malaria Anaemia
SNPs	Single Nucleotide Polymorphisms
SSA	Sub-Saharan Africa
STEVOR	Subtelomeric Variable Open Reading frame
sWGA	Selective Whole Genome Amplification
TCS	Transitive Consistency Score
TNF-α	Tumor Necrosis Factor Alpha
TRAP	Thrombospondin-related Anonymous Protein
VAR2CSA	Variant Surface Antigen-2 Chondroitin Sulphate 1
VCF	Variant Call Format
VSA	Variant Surface Antigen
WHO	World Health Organization

CHAPTER 1

INTRODUCTION

1.1 Background

Malaria is a disease caused by infection with *Plasmodium* species and it remains one of the world's most significant public health problems. An estimated 249 million cases of malaria were reported worldwide in 2022, with 608,000 deaths. Sub-Saharan Africa (SSA) accounted for 93.6 % of all cases and deaths, with pregnant women and children below five years being the most vulnerable groups (WHO, 2023). Recent world malaria reports have either indicated a stall or an increase in global malaria incidence and deaths, especially in Africa. This is partly attributed to the mosquito and parasite's remarkable genetic variation, and resistance to insecticides and antimalarial drugs (Ashley & Phyo, 2018; Ndiath, 2019; Zhu et al., 2022), or due to the disruption of health services during the COVID-19 pandemic. (WHO, 2021, 2022a). *Plasmodium* parasites generally spread to people through the bite of female *Anopheles* mosquito species. Six *Plasmodium* species; *P. vivax*, *P. ovale, P. malariae, P. falciparum, P. knowlesi*, and *P. cynomolgi* infect humans among which *P. falciparum* is the major cause of malaria disease and death (Foko et al., 2022; Jeyaprakasam et al., 2020; WHO, 2019).

The disease's clinical manifestations are brought on by the parasite during its erythrocytic stage of development. This stage begins when merozoites released from the liver invade the erythrocytes. Erythrocyte invasion by the merozote involves the interaction of many parasite ligands with red blood cell (RBC) surface receptors (Cowman et al., 2017). Two known parasite protein families: the erythrocyte binding antigens (EBA) protein family (EBA-175, EBA-181, EBL-1, and EBA-140) and the reticulocyte binding-like homologue (Rh) protein family are major determinants of merozoite invasion (Beeson et al., 2016; Tham et al., 2012). These parasite proteins or ligands bind to different red blood cells (RBCs) receptors some of

which (the glycophorins) contain a nine-carbon alpha-keto acid sugar backbone known as sialic acid while others (Complement receptor 1, Band 3, and Basigin) do not. During erythrocyte invasion, the EBA protein family members such as EBA-175, EBL-1, and EBA-140 bind to glycophorins (GP) A, B, and C respectively (Adams et al., 2001; Ashline et al., 2015; Mayer et al., 2009), while the Rh family members, Rh4 and Rh5, bind to complement receptor 1 (CR1) and basigin respectively (Crosnier et al., 2011; Tham et al., 2010). Hence, parasites that use the EBA protein family members to invade the erythrocyte are known to be sialic acid (SA) dependent while those that make use of the Rh family members are sialic acid independent parasites (Ord et al., 2012). This allows the parasite to catalyse invasion through a multitude of different ligand-receptor interactions, resulting in a diversity of invasion pathways (Hadley et al., 1987). Differential expression of invasion ligand genes can result in the utilisation of various receptors during erythrocyte invasion and hence, different invasion pathways. Likewise, polymorphisms in both invasion ligands and erythrocyte receptors can also result in variability in invasion pathways and disease outcomes (Iriko et al., 2008; Verra et al., 2006). Previous studies in The Gambia over a decade ago showed the dominance use of sialic acid-dependent invasion pathways when transmission was comparatively higher (Baum et al., 2003; Gomez-Escobar et al., 2010). However, over the years, malaria transmission in The Gambia have drastically reduced (from 4 % in 2010 to 0.2 % in 2022) (WHO, 2022a). Hence, it is unknown how erythrocyte invasion mechanisms and gene expression profile have been impacted by this shift. Additionally, with the current global goal of malaria elimination through genomic surveillance, increased research, and intensified interventions/control strategies, such as the roll-out of malaria vaccines (e.g RTS'S/R21), the host-parasite interaction essential for erythrocyte invasion might have devolope new adaptive mechanisms including switching of dominant invasion pathways, changes in ligand gene expression or sequence polymorphisms in both invasion ligands and erythrocyte receptors. Therefore, knowledge of the variety of

erythrocyte invasion pathways used by extant African *P. falciparum* parasites, the related transcriptional heterogeneity, and genetic variation is essential for future strategies such as new drugs and vaccines targeting the erythrocytic stage of malaria parasite life cycle.

Furthermore, the human genome is under significant selective pressure from malaria (Kwiatkowski, 2005) hence, different variants at host receptors might have arisen because of selection by specific *P. falciparum* strains, with specific receptor-ligand allelic combinations (e.g., the erythrocyte binding proteins and the glycophorins) influencing parasite invasion, growth, and disease differently. However, in most previous studies, exploring the association between genetic variants in parasites or humans and disease phenotypes have mostly not considered the co-selection of variants in humans and parasites. Additionally, continuous molecular surveillance has been recommended (WHO, 2022a) as one of the malaria elimination strategies and since the parasite is not independent of the host, dual studies on both humans and parasites could identify dual signatures of selection among malaria-positive individuals necessary for the development of control strategies.

Moreover, several structural variants (duplications, deletions, and insertions) have been identified in the human genome, especially in the glycophorin (GP) locus. Some of these structural variants like the duplications of glycophorin A (GPA) and glycophorin B (GPB), referred to as duplication 4 (DUP4) confer some level of protection against severe malaria in East Africa (Leffler et al., 2017; Louzada et al., 2020; MalariaGEN, 2015). However, other common variants in West Africa like the GPB deletion 1 (DEL 1) and deletion 2 (DEL 2) are not well characterised and their relationship with malaria is yet to be established.

In this study, parasite culture, targeted transcriptomics, and next-generation sequencing approaches were used to interrogate genome (parasite)-genome (host) interactions and provide a better understanding of *P. falciparum* invasion mechanisms, gene expression profile, and genetic variations in both human receptors and parasite ligands in The Gambia.

1.2 Objectives, Hypotheses, and Rationale

1.2.1 Overall Objective

To explore host-parasite genetic and transcriptomic correlates of *P. falciparum* erythrocyte invasion phenotypes, and malaria outcomes in The Gambia

1.2.2 Hypothesis 1

Erythrocyte invasion pathways of parasites in The Gambia will be dominated by sialic-acid independent mechanisms driven by the expression of specific ligands

1.2.3 Objective 1

To determine the genetic variations and transcriptional changes of *P. falciparum* invasion ligands in isolates with different invasion phenotypes in The Gambia

1.2.4 Rationale 1

The malaria parasite has a redundancy of ligands that can be used for erythrocyte invasion. The switching of parasite lines between invasion pathways may be determined by several factors. First, parasites express different invasion ligand genes as a response to the environment created by the immune system or control interventions. Different invasion ligands bind to different erythrocyte receptors resulting in alternative invasion pathways. Secondly, sequence polymorphisms in parasite ligand (or receptor) genes may result in variation in efficacy or use of specific invasion pathways. Prior research conducted in The Gambia more than ten years ago revealed the dominance of sialic acid-dependent invasion pathways at a time when malaria transmission was comparatively higher (Baum et al., 2003; Gomez-Escobar et al., 2010). However, over the last 15 years, malaria transmission in The Gambia has drastically reduced (from 4 % in 2010 to 0.2 % in 2022), and The Gambia is currently in the pre-elimination phase of malaria (WHO, 2022a). Hence, it is unknown how erythrocyte invasion mechanisms have

been impacted by this shift from high to low malaria transmission. This objective is therefore powered by a multi-prong approach: cell biology of invasion, targeted transcriptomics, and PCR-based genotyping approaches to understand the invasion pathways, genetic diversity, and transcriptional variation of *P. falciparum* from malaria positive individuals (1 – 70 years old) in The Gambia.

1.2.5 Hypothesis 2

Specific *P.falciparum* invasion ligand and human receptor interactions may result in adaptive variants that drive infection rates

1.2.6 Objective 2

To determine genetic variation in *P. falciparum* invasion ligands and their cognate human receptors in malaria cases from The Gambia

1.2.7 Rationale 2

The malaria parasite invasion and replication in red blood cells (the most common blood cell in humans) is responsible for most of the symptoms of the disease. This has over time exerted enormous selective pressure on the human population, leaving malaria-associated polymorphisms and signatures in the genome (Kwiatkowski, 2005). Thus variations in receptors such as the glycophorins and several blood group factors are associated with protection (Leffler et al., 2017). Although variations in invasion ligands and erythrocyte receptors have been extensively studied separately, the distribution of alleles within the same infected individuals has received far less attention. Thus, to generate baseline data on host-paraste interactions from the same study/individuals in The Gambia, it is important to start with the most common erythrocyte receptors, parasite ligands and surface proteins like GPA, GPB, GPC, CR1, EBA-175, EBA-181, EBA-140, EBL-1, Rh4, Rh5, MSP1, MSP6, DBLMSP, CLAG2, CLAG8, and SURFIN4.2. This could be useful in vaccine and drug design, development of control strategies as well as enhancing understanding of host-parasite

interactions, disease development, and progression. Additionally, continuous molecular surveillance has been recommended (WHO, 2022a) as one of the malaria elimination strategies and since the parasite is not independent of the host, dual studies on both humans and parasites to identify cognate signatures of selection among malaria-positive individuals of all age groups will help in the design of more effective control interventions.

1.2.8 Hypothesis 3

Glycophorin B deletion variants (Deletion 1 and Deletion 2) in The Gambia vary across ethnic groups and local malaria transmission patterns.

1.2.9 Objective 3

To determine the distribution of glycophorin B receptor variants (Deletion 1 and Deletion 2) among individuals with clinical malaria in The Gambia

1.2.10 Rationale 3

The human glycophorins are crucial elements of *P. falciparum's* erythrocytic cycle because of their function as surface receptors for erythrocyte invasion. Several structural variants (deletions, duplications, and insertions), some protective against severe malaria, have been identified in the glycophorin locus across African human populations (Leffler, et al., 2017). However, only a few of these variants like duplication 4 (Dup 4) commonly called the Dantu has been thoroughly investigated and functionally identified as being connected to severe malaria in East Africa (Kariuki et al., 2020; Louzada et al., 2020). Due to high sequence homology between the glycophorins, it has been challenging to develop specific PCR-based genotyping assays for effective characterisation of other common variants like the glycophorin B deletions. Currently, eight glycophorin B deletions (GPB DEL 1 to 8) have been identified among which deletions 1 and 2 are highly prevalent in West Africa. The association of these deletions especially the two most common ones in West Africa with malaria is not known and their distribution in each West African country has not been determined. Fortunately, recent

interest in GPB deletions has resulted in the development of PCR-based genotyping assays (Amuzu et al., 2021; Lane et al., 2020) enabling a cheaper and faster-targeted approach to studying the distribution of these deletions and their association with malaria. Given that host genetic variations have been well documented to influence malaria outcomes (Kwiatkowski, 2005) and malaria prevalence in The Gambia is very low (0.2 %), this objective takes advantage of existing PCR-based assays to determine the distribution of the two most common GPB deletions (Deletion 1 and 2) among malaria individuals of all age groups (1 to 70 years) in The Gambia.

1.3 Significance of the Study

The findings of this study are essential for enhancing genomic surveillance as The Gambia intensifies malaria elimination efforts. By uncovering how malaria parasite strains use different mechanisms to invade the erythrocytes and how these affect disease dynamics, this research could guide new control strategies. Moreover, it has provided valuable insights into the parasite biology, malaria pathogenesis, and the complex parasite's life cycle, while also identifying single nucleotide polymorphisms (SNPs) in both the parasite and host that could serve as potential drug and vaccine targets. The genetic data generated in this study have improve our understanding of malaria severity through the identification of SNPs associated with malaria susceptibility. This could also go a long way to inform better treatments and interventions, and contribute to developing novel diagnostics and vaccines.

CHAPTER 2

LITERATURE REVIEW

2.1 Malaria Disease and Causative Parasite Species

The bite of an infected female Anopheles mosquito causes the vector-borne infectious disease malaria, which is spread to vertebrates by a unicellular eukaryotic protozoan of the genus *Plasmodium*. Only six of the over 250 *Plasmodium* species are known to infect humans (Leclerc et al., 2004): *P. ovale, P. vivax, P. malariae, P. falciparum, P. knowlesi*, and *P. cynomolgi* (Foko et al., 2022; Jeyaprakasam et al., 2020; WHO, 2019). Majority of moderate and severe cases of the disease are caused by *P. falciparum*, which is most common in Sub-Saharan Africa (SSA). On the other hand, *P. vivax* is a far more widespread disease in Asia, Latin America, and some regions of Africa because it can grow in the Anopheles mosquito (its vector), at higher altitudes and lower temperatures (Battle & Baird, 2021). The two less common species *P. ovale* and *P. malariae* are more prevalent in Africa and worldwide respectively. Two closely related but separate species of *P. ovale* have been identified: *P. ovale wallikeri* and *P. ovale curtisi* (Fuehrer & Noedl, 2014; Sutherland et al., 2010). *P. knowlesi* and *P. cynomolgi* are the only two zoonotic *Plasmodium* species known to affect both humans and monkeys and are both prevalent in Malaysia, and South-East Asia. While *P. knowlesi* has been well-established to infect humans, *P. cynnomolgi* was just recently reported by studies in Malaysia to infect humans (Hartmeyer et al., 2019; Jeyaprakasam et al., 2020).

The female Anopheles mosquito maintains the circulation of all six *Plasmodium* species in populations. Over 41 Anopheles species are recognised to be the primary vectors of malaria transmission to humans, including those of the Gambiae complex, (Coetzee et al., 2013; Coetzee & Fontenille, 2004), Funestus (Cohuet et al., 2003), Nili (Kengne et al., 2003), and Moucheti (Kengne et al., 2007). Among these, *An. gambiae*, *An. colluzzi, An. nili, An.*

moucheti, An. funestus, An arabiensis, and *An melas* are the main species in SSA (Longo-Pendy et al., 2022). Successful parasite transmission to the human host depends on the number of parasites in circulation (parasite abundance), parasite survival after a blood meal (longevity), parasite carriage in the salivary glands (capacity), and nearness of humans to breeding sites (contact with humans) (Massey et al., 2016).

Malaria predominantly affects tropical and subtropical regions worldwide and it is usually asymptomatic or characterised by symptoms like fever associated with joint pains, chills, headaches, myalgia, and malaise which can either result in mild (mild or uncomplicated malaria) or life-threatening complications (severe malaria) (Boushab et al., 2020; Yusuph et al., 2019). Several strategies are being used to control malaria including vector control strategies like the destruction of breeding grounds, insecticide sprays, and the use of long-lasting insecticidal nets (LLINs). Drugs have also been used for treatment and prophylaxis at individual and mass scales to curb morbidity and mortality due to the disease. More recent additions to this arsenal are vaccines, the RTS,S, and R21/Matrix-M, which have been approved for mass rollout by the WHO (WHO, 2023). Despite these control strategies, the disease persists and devastates the tropics and subtropics, especially in Sub-Saharan Africa to date.

2.1.1 Global Burden and Distribution of Malaria

The 2023 world malaria report estimates that in 2022, there were 249 million cases of malaria in 85 countries where the disease is endemic (Figure 2.1) compared to 241 million cases in 2020 and 247 million in 2021 (WHO, 2021, 2022a, 2023). Although there has been a rise in malaria cases since 2016, the first year of the COVID-19 pandemic (2020) saw the largest yearly increase of around 13 million cases, and COVID-19 disruptions resulted in an additional 13.4 million cases between 2019 and 2021. Likewise, malaria deaths have also increased from 409,000 cases in 2019 to 631,000 in 2020, 610,000 cases in 2021, and 608,000 cases in 2022

leading to a projected rise in mortality of 63,000 people between 2019 and 2021, which is linked to the COVID-19 pandemic's disruption of essential malaria services (WHO, 2020, 2021, 2022a, 2023). The WHO African region has the highest disease burden. Globally, 93.6 % of malaria cases and 95.4 percent of deaths in 2022 were reported to have occurred in the region. Also, the WHO South-East Asia region was home to 2 % of the global malaria cases in 2022 with India accounting for 65.7 % of these cases (WHO, 2023).

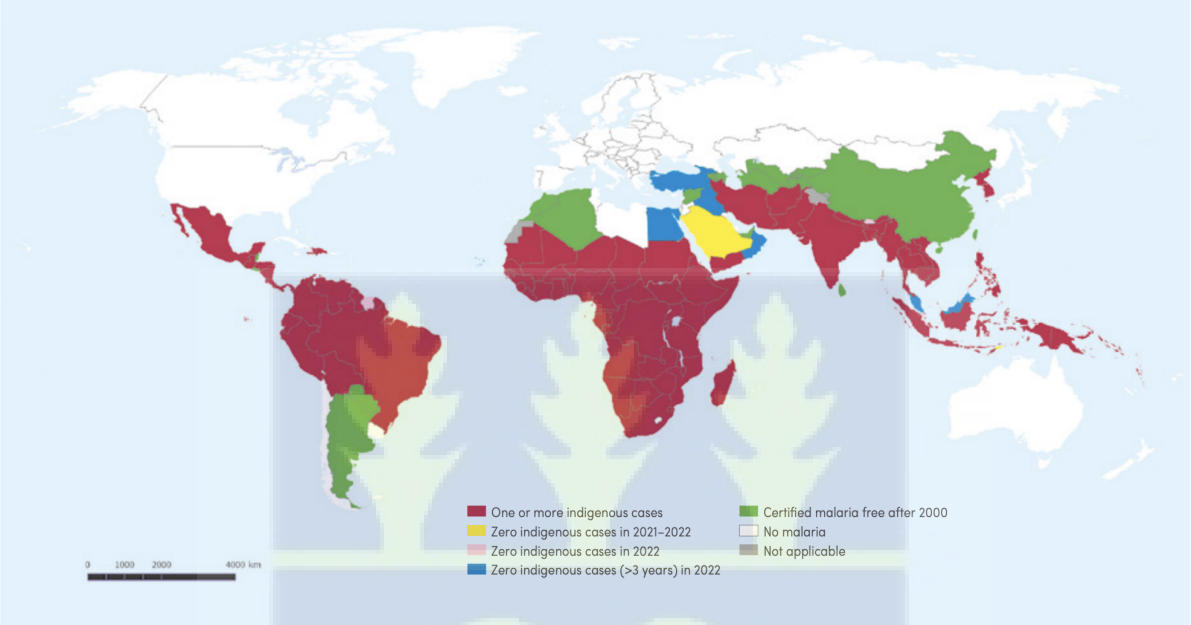


Figure 2.1: The Status of countries with native cases of malaria in 2000 and by 2022 (Source: WHO,2023)

2.1.2 Malaria Status in The Gambia

The Gambia is a small country (10,689 square kilometres) in Western Africa, fronted by a 60 km atlantic ocean and surrounded on three sides by Senegal. It is home to 2,848,583 million people with a density of 176 people per square kilometre (Gambia population: https://www.worldometers.info/world-population/gambia-population/). The urban and peri-urban areas are home to about 57 % of the population. Administratively, it is divided into six regions: West Coast, North Bank, Central River, Lower River, Upper River, and Great Banjul Regions (regions of The Gamia, https://www.mappr.co/counties/gambia/).

A brief wet season (June to October) and a lengthy dry season (November to May) define the climate. The wet season creates breeding sites for mosquitoes, especially three species of *Anopheles gambiae* complex (*An. gambia s.s, An. arabiensis* and *An. melas*), mostly responsible for all malaria transmission in the country (Bøgh et al., 2003). Almost all malaria cases are due to *P. falciparum* with annual peaks between September and November. Significant advancements have been achieved in the battle against malaria over the years with a continuous decline in malaria incidence nationwide (Ceesay et al., 2008). In 2017, malaria prevalence in The Gambia was 0.2 % compared to 4 % in 2011 (WHO Regional Office for Africa, 2017). Malaria cases gradually reduced from 361.5 cases per 100,000 in 2001 to 87.3 cases per 100,000 in 2020 and 2021. One of the first five nations to achieve the 40 % decrease in case incidence required by the global technical strategy (GTS) for malaria was The Gambia. Because of this, it is expected to meet the GTS targets by 2025 and 2030 and has refocused its activities on the subnational elimination of malaria.

2.2 Life Cycle of *Plasmodium falciparum*

The *P. falciparum* life cycle consists of three different stages of development (the sporogonic, the pre-erythrocytic, and the erythrocytic) in two different hosts (human and mosquito vector). The pre-erythrocytic and erythrocytic cycles take place in the liver and erythrocytes respectively, whereas the sporogonic cycle occurs in the mosquito's digestive system (Figure 2.2). During a blood meal, an uninfected individual becomes infected when sporozoites, the infectious form of the parasite, are injected into their skin by an infected female Anopheles mosquito. The motile inoculated sporozoites move via several kupffer cells to invade the hepatocytes in the liver (Mota et al., 2001). This hepatocyte invasion is mediated by the *P. falciparum* circumsporozoite protein (*PfCSP*), the major surface protein of the sporozoite (Ménard et al., 1997). The sporozoites developed into tissue schizonts (a structure that contains merozoites, another stage of the parasite). When the tissue schizont reaches maturity, it bursts

and discharges merozoites into circulation. This first developmental step of the parasite is called the pre-erythrocytic cycle and it takes about six to ten days during which the infected person remains asymptomatic (Tarun et al., 2006).

The erythrocytic stage of the *Plasmodium's* life cycle begins when merozoites are released into the bloodstream (Sturm et al., 2006). Merozoites proliferate and transform into several morphological phases in erythrocytes, beginning with the ring, followed by the trophozoite, and then the mature schizont. The mature schizont ruptures, releasing several merozoites into circulation each of which can reinitiate the erythrocytic cycle in 30 seconds by invading a fresh erythrocyte. This erythrocytic phase of the parasite's life cycle, which causes the clinical signs of malaria, takes 48 hours to complete. However, not all erythrocytic parasites grow into schizonts; some of them differentiate into gametocytes in circulation (sexual erythrocytic stage). Two processes lead to gametogenesis: a proportion of the merozoites from the hepatocyte directly become committed to developing into gametocytes, bypassing the erythrocytic schizogony stage (Bancells et al., 2018) while others experience erythrocytic schizogony, a stage in which every committed schizont's merozoite grow into either female or male gamete, but never both sex (Poran et al., 2017; Silvestrini et al., 2000). Gametogenesis occurs because of the de-repression of an epigenetically silenced *ap2-g* gene locus in the replicating parasite (Bechtsi & Waters, 2017; Kafsack et al., 2014). The committed gametocyte undergoes a series of morphological developmental stages (I-V) wherein, stages I-IV (immature stages) sequester in the bone marrow to evade the host immunity while stage V (mature infectious stage) is released in circulation to facilitate mosquito ingestion (Nixon, 2016). These transmissible stage V gametocytes are ingested by the female Anopheles mosquito during a blood meal. The male gametocyte ex-flagellate is activated by pH changes in the mosquitoes' midgut after ingestion, resulting in the production of eight microgametes that are needed to fertilise the macrogametes produced by the female gametocyte (Janse et al.,

1988; Sinden et al., 2010). After fertilisation, a diploid zygote is produced in the mosquito's midgut, which then goes through DNA replication to become an ookinete, an oocyst, and eventually haploid sporozoites that can migrate to the salivary gland and infect the human host again during subsequent blood meals (Figure 2.2).

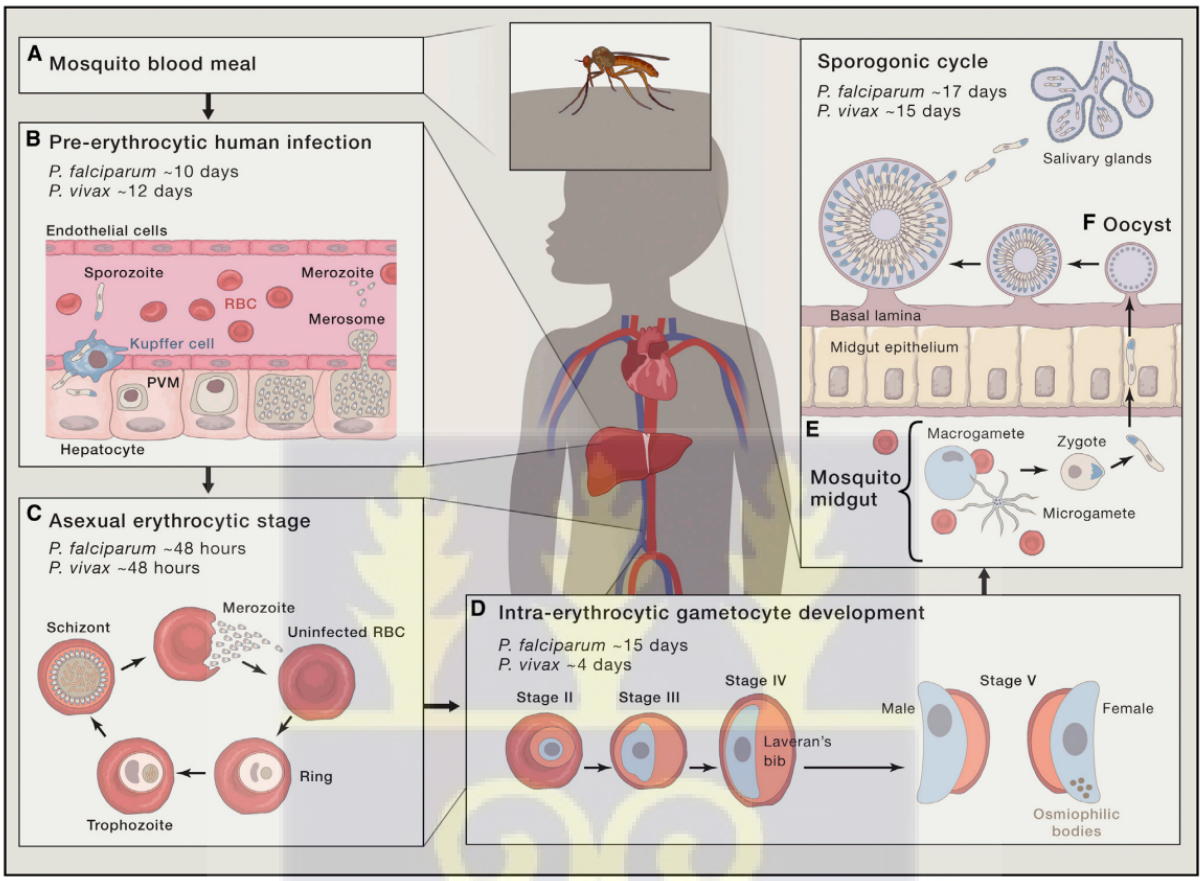


Figure 2.2: The *P. falciparum* life cycle (Source: Cowman et al., 2016).

RBC = Red blood cells and PVM = Parasitophorous vacuolar membrane. During a blood meal, sporozoites are injected into the host by an infected female Anopheles mosquito (A). Sporozoites infect the hepatocyte, and develop into mature schizonts which rupture and release merozoites B). Merozoites infect the erythrocyte and develop into rings, trophozoites, and mature schizonts which rupture releasing merozoites (C). A subset of the ring stage differentiates into gametocytes which develop from immature stages (I-IV) to mature stage V gametocytes that are ingested by the mosquito vector while blood feeding (D). The micro and macrogametes of the female and male gametocyte fuse to produce a zygote in the mosquitoes'

midgut (E). When the zygote matures into an oocyst, it bursts and releases motile sporozoites, which go to the salivary glands in preparation for injection into the next human host during another blood meal (F).

2.3 The *Plasmodium* Genome

In 2002, the *P. falciparum* genome sequencing consortium sequenced the genome of *P. falciparum* from the 3D7 strain for the first time (Gardner et al., 2002). The nuclear genome of 3D7 is 23.3 Mb in size with a GC (Guanine and Cytosine) content of 19.4 % and an AT (Adenine and Thymine) content of 80.6 %. This high AT content makes the genome distinct from other eukaryotes and makes it unstable as well as the DNA difficult to handle. It is made up of 5542 annotated protein-encoding genes organised into 14 linear chromosomes with sizes ranging from 0.643 Mb to 3.29 Mb. Majority of the genes range between 2000 to 2500 bp with the majority of highly amplified and hypervariable gene families in the chromosome's sub-telomeric regions (Rovira-Graells et al., 2012). The sub-telomeric region ranges from 60 kb to 120 kb and has a well-studied structure constituting one to six telomere-associated repeat elements, several surface antigens, and virulence genes known to play crucial roles in virulence, antigenic variation, and immune evasion (De Bruin et al., 1994).

P. falciparum possesses two genomes in addition to its nuclear genome: an apicoplast genome (35 kb) and a mitochondrial genome (6 kb). Most of the mitochondrial DNA is composed of tandem repeats with short sequences, and it has only three genes that code for three different proteins (Cytb, Cox3, and Cox1) (Hikosaka et al., 2011). Moreover, the mitochondrial genome does not encode for any transfer RNA (tRNA) hence, export of tRNA occurs from the cytosol. *P. falciparum's* apicoplast genome encodes 30 proteins involved in important pathways like the haeme metabolism, fatty acid, or isoprenoid synthesis pathway (Foth & McFadden, 2003). The five blocks (SB 1-5) that make up the very complex sub-telomeric portions of the *P. falciparum* genome are oriented and composed in a way that encourages mitotic recombination

between telomeres (Gardner et al., 2002). The chromosomes are made up of conserved and polymorphic domains with the conserved domains located within the core regions where the housekeeping genes are dominants. The polymorphic domains are in the sub-telomeric blocks and are AT-rich. They play potential roles in gene conversion, chromosome pairing, and recombination (De Bruin et al., 1994). To date, the 3D7 genome is continually being revised and used as a reference.

2.4 Pathogenesis and Pathophysiology of *P. falciparum* Malaria

2.4.1 Clinical Outcomes of Malaria

Malaria is classified as either uncomplicated (mild) or complicated (severe) based on its clinical consequences. Uncomplicated malaria can either be asymptomatic or symptomatic whereby, the mild symptoms mostly manifest in three stages: the cold stage which consists of a sensation of cold and shivering, the hot stage where the patient suffers from headaches, vomiting, fever, and seizures in young children (CDC, 2022). When mature schizonts rupture in an infected erythrocyte, poisonous bilirubin and hemozoin are released, causing these clinical signs to manifest. A complete dose of appropriate malaria drugs will clear the infection, while the return of symptoms after treatment is either due to new infection, resistance to the drugs used, or incomplete treatment. However, in some untreated or partially treated individuals (mostly children), the immune system is unable to control the initial infection, leading to the progression to severe or complicated forms of malaria. When crucial organ failures or irregularities in the blood are observed in patients, they are said to have severe malaria, either clinically or by laboratory testing. Severe malaria is mainly associated with acute renal failure, severe anaemia, acute respiratory distress syndrome (ARDS), cerebral malaria, haemoglobinuria, abnormal blood coagulation, low blood pressure, metabolic acidosis, hypoglycaemia, and hyperparasitemia (more than 5 % of red blood cells infected) (CDC, 2022). Major predisposing factors of severe malaria are age (>65 years), naïve or lack of immunity

(children under five or non-immune adults from industrialised countries returning from malaria endemic zones), pregnancy, malaria co-occurrence with other health problems, delay in medical treatment, and the absence of prophylaxis (Bruneel et al., 2003; Schwartz et al., 2001).

2.4.1.1 Severe Malaria Anaemia (SMA)

In endemic nations or regions with high malaria transmission rates, SMA is the most frequent clinical consequence of severe malaria and the main reason why children are hospitalised. It is generally characterised by haemoglobin level below 5 g/dl (haematocrit lower than 15 %) for children with parasitaemia and less than 7 g/dl (haematocrit of less than 20 %) for adults (WHO, 2022a). Some characteristics of SMA are lysis of infected and uninfected erythrocyte (Price et al., 2001), erythrocyte sequestration in the spleen (Buffet et al., 2009), suppressive bone marrow, chronic malaria transmission in holoendemic areas, and decreased and/or inefficient erythropoiesis (Phillips et al., 1986). When haemolysis is caused by parasites and/or anti-malarial medications which prevent the proper synthesis of erythrocytes, the body is unable to restore the erythrocyte pool that has been depleted. Prior research in Gambian children has shown that erythroid hyperplasia with dyserythropoiesis is the hallmark of SMA (Abdalla & Pasvol, 2004).

2.4.1.2 Cerebral Malaria (CM)

Cerebral malaria is the highest dramatic and complicated type of severe malaria and the main cause of long-term neurocognitive impairments. Despite appropriate therapy, the mortality is 15 % among individuals affected and about 30 % of those who survived have neurologic sequelae (Postels et al., 2012). Patients with CM usually suffer from multiple seizures and loss of consciousness resulting in confusion and fatal coma. Other clinical features of CM include repeated convulsions (≥ two episodes in 24 hours), prostration, pulmonary oedema, respiratory distress (acidotic breathing), and deep breathing. CM is usually associated with the production of excess tumour necrosis factor-alpha (TNF-α) resulting from obstruction of the cerebral

microvasculature. The obstruction of the cerebral microvasculature emerges as a result of erythrocyte sequestration in the brain resulting in inflammation, bleeding, and swelling (Adams et al., 2002), and rosette formation (Carlson et al., 1990; Treutiger et al., 1992). During the erythrocytic cycle, the ring stage develops into trophozoites which cling to epithelial cells in the microvascular circulatory system as a means of evading the immune system (Autino et al., 2012). The *P. falciparum* erythrocyte membrane protein-1 (PfEMP-1) mediates this process. PfEMP-1 is encoded by the *var* genes and is often expressed on the surface of erythrocytes infected with trophozoites and schizonts (Hviid & Jensen, 2015). Moreover, the PfEMP-1 extracellular component consists of several adhesive domains that bind to receptors like intracellular adhesion molecule-1 (ICAM-1) and endothelial progenitor cell receptor (EPCR) (Ochola et al., 2011; Turner et al., 2013), causing vaso-occlusion and inflammation. Notably, human CM without proof of parasite sequestration in the cerebral tissues has equally been documented (De Souza et al., 2010; Milner, 2010). About a quarter of children who recover from CM usually experience permanent neurological defects within three to seven years.

2.4.2 Erythrocyte Invasion

2.4.2.1 The Merozoite and its Secretary Organelles

The parasite's erythrocytic cycle usually begins when the merozoite (invading parasite stage) invades the erythrocyte. The merozoite is an egg-shaped invasive form of the *Plasmodium* parasite. It is 1-2 μm in length and its primary goal is to invade the erythrocyte and ensure continuation of the erythrocytic cycle. Hence, the structure of the merozoite is purposely designed to execute this goal (Cowman et al., 2017). The merozoite's inner side is known as the inner membrane complex which houses the internal organelles while the mid-section constitutes the nucleus, mitochondrion, and the plastid (Preiser et al., 2000). The cytoskeleton is located in the cortical space, which is a gap between the inner membrane complex and the outer plasma membrane (Preiser et al., 2000). The merozoite also contains different organelles

like micronemes, rhoptries, dense granules, and exonemes which facilitate the invasion process upon contact with the erythrocyte (Figure 2.3). Adhesins released by micronemes facilitate the binding of the merozoite to erythrocytes whereas the rhoptries are responsible for the invasion process (Healer et al., 2002; Kremer et al., 2013). The exonemes release the enzyme subtilisin-1 (SUB1) into the parasitophorous vacuole to enhance merozoite egress. Additionally, the dense granules of the merozoite are typically secreted during various phases of the erythrocytic cycle and are essential for the release of certain ligands at every stage of the invasion process (Cowman et al., 2017; Yeoh et al., 2007).

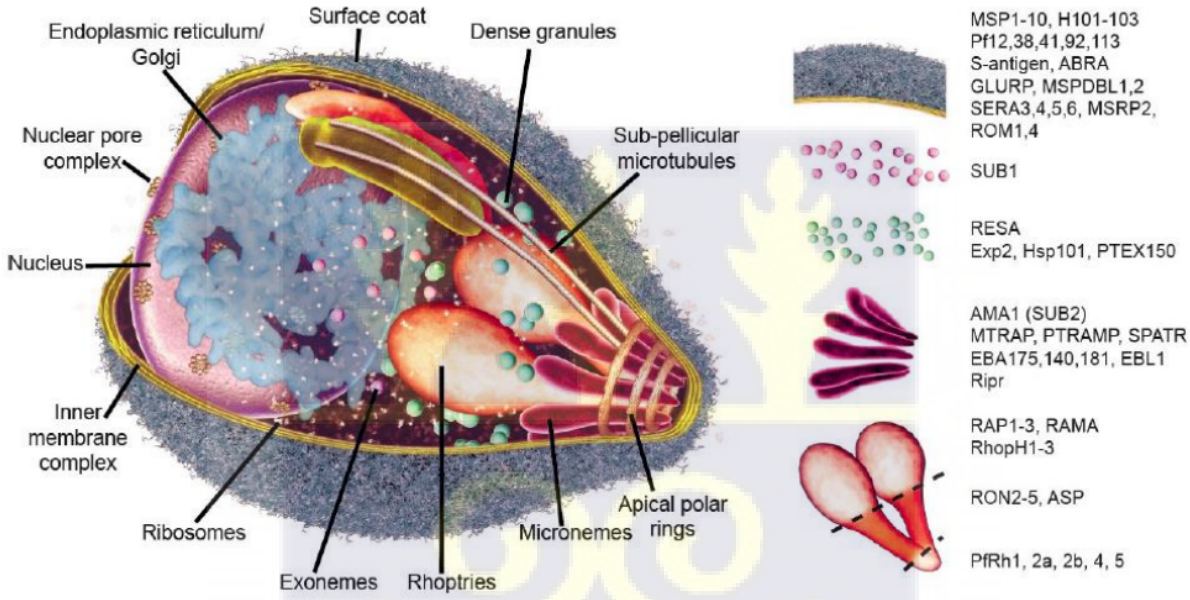


Figure 2.3: Structure of the merozoite and its secretory organelles (Source: Cowman et al., 2012)

2.4.2.2 The Molecular Basis of Merozoite Invasion

Merozoite invasion is a well-organized process involving a cascade of molecular interactions between specific parasite ligands and erythrocyte receptors through four distinct steps (initial attachment, merozoite reorientation, tight junction formation, and active invasion). Attachment of the merozoite to the erythrocyte is the first step of invasion, mediated by the merozoite

surface proteins (MSPs) (Figure 2.4A). Once the merozoite is attached, it reorientates itself thereby bringing its apical poles in contact with the erythrocyte membrane enabling firm attachment to the erythrocyte. This step is mediated by the erythrocyte-binding antigens (EBA) and reticulocyte-binding homologues (Rh) protein families (Figure 2.4B). Following these interactions, the PfRh5 complex binds to its receptor basigin, leading to micronemes secretion and calcium influx to the erythrocyte to initiate downstream invasion processes (Figure 2.4C). The rhoptry neck proteins (RONs) complex is then released onto the surface of the erythrocyte to interact with the apical membrane antigen-1 (AMA-1) (Figure 2.4D) and mobilise the merozoite internal actinomyosin complex, enabling the merozoite to push itself into the erythrocyte. Coincidentally, the rhoptry proteins are released from the merozoite, resulting in the parasitophorous vacuole's development which encloses the merozoite (Figure 2.4E-G). Echinocytosis and erythrocyte recovery come next (Figure 2.4H and I) (Cowman et al., 2016; Weiss et al., 2016).

Since the erythrocyte does not have a nucleus or other organelle, *Plasmodium* species use their invasion machinery and express the necessary invasion-related proteins. The process is closely coordinated, involving more than 400 proteins (Hu et al., 2010) and it takes less than 30 seconds (Treeck et al., 2009).

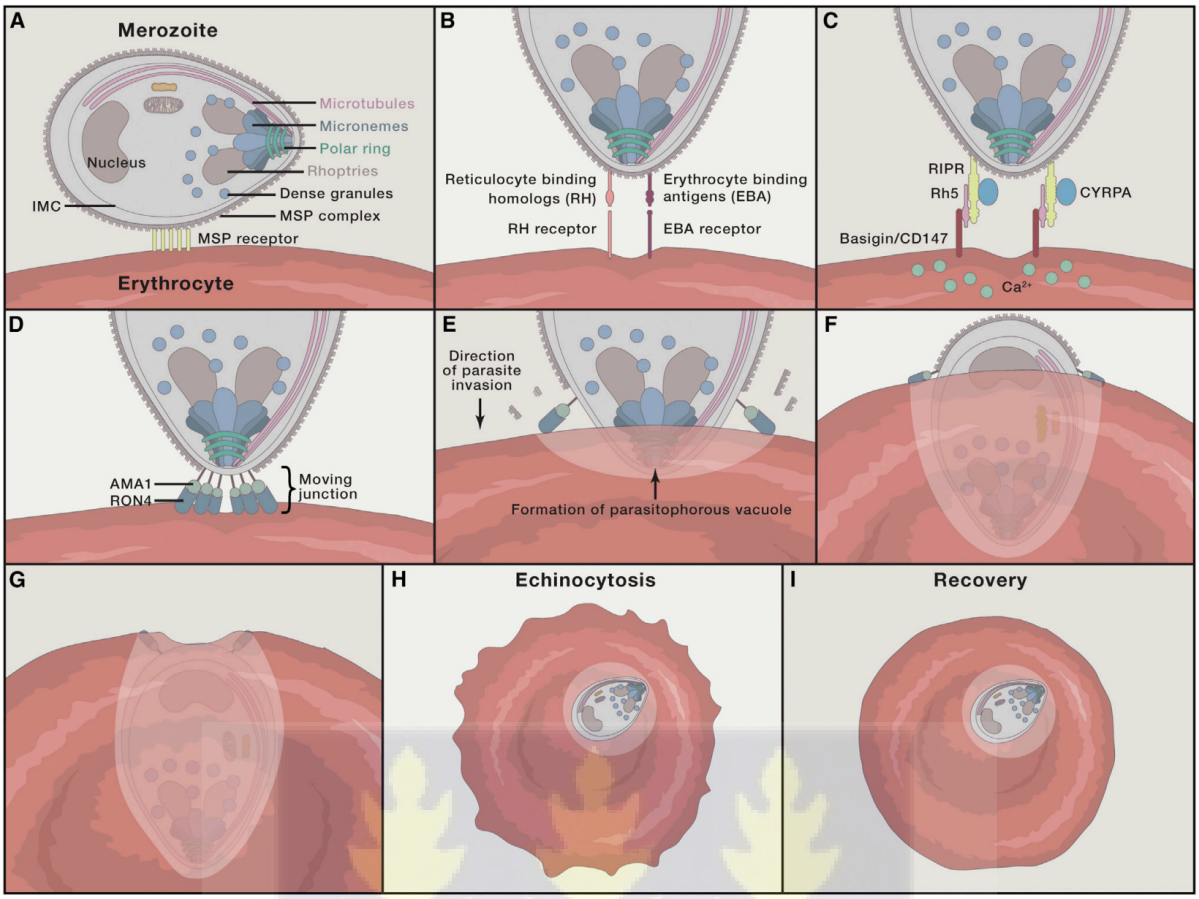


Figure 2.4: The merozoite and its invasion mechanism (Source: Cowman et al., 2016).

IMC = Inner membrane complex, MSP = Merozoite surface protein, RH = Reticulocyte binding homolog, Rh5 = Reticulocyte binding homolog 5, AMA1 = Apical membrane antigen 1, RON4 = Rhoptry Neck Protein 4, RIPR = Reticulocyte binding protein homolog 5 interacting protein, CYRPA = Cysteine-rich protector antigen, CD147 = Cluster of differentiation 147, and EBA = Erythrocyte binding antigen

2.4.2.2.1 The Merozoite's Attachment

Initial contact is the first step of invasion whereby, new merozoites are released from ruptured schizonts into the bloodstream. The released merozoites identify uninfected erythrocytes and adhere to their plasma membrane in a reversible manner leading to a minor deformation of the erythrocyte membrane (Weiss et al., 2016). The surface of the merozoite is comprised of glycosylphosphatidylinositol (GPI)-anchored proteins which are mainly members of the

merozoite surface protein (MSP) family (Sanders et al., 2005). The MSPs (MSP-1 to MSP-10) mediate the initial contact of the merozoite to the erythrocyte (Lin et al., 2014) (Figure 2.5). MSP-1 which is the most common MSP is 190 to 200 kDa and is made up of four fragments with molecular weights of 83, 30, 38, and 42 kDa. All MSP-1 fragments—aside from the 19-kDa fragment are eliminated from the merozoite's surface during initial contact and parasite entry (Dluzewski et al., 2008). Heparin-like molecules found on the erythrocyte surface and involved in parasite invasion bind to MSP-1_{42} and MSP-1_{33} fragments to facilitate the invasion process. One major erythrocyte transmembrane protein glycophorin A (GPA) has also been shown to interact with MSP-1_{42} and two extracellular regions of Band 3 during initial attachment (Baldwin et al., 2015; Goel et al., 2003) (Figure 2.5). About 25 % of the erythrocyte membrane is made up of Band 3, a 110-KDa glycoprotein and it is an important receptor involved in the sialic acid-independent invasion pathways of *P. falciparum*. It associates with the GPA receptor and increases the possibility of attachment of the merozoite ligands to the erythrocytes (Baldwin et al., 2015).

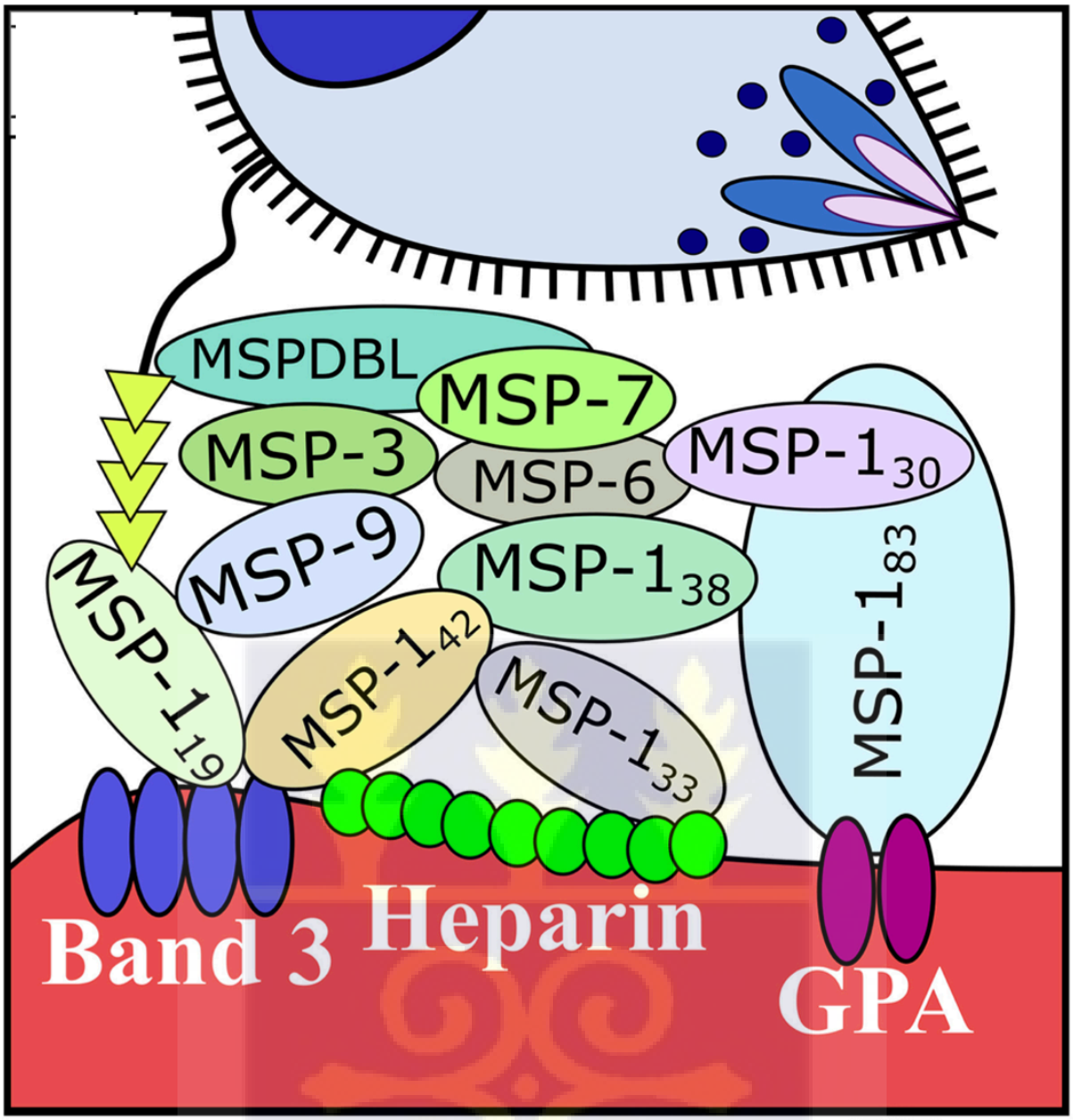


Figure 2.5: Molecular interaction between MSPs and erythrocyte receptors during the initial stage of invasion (Source: Molina-Franky et al., 2022). MSP = Merozoite Surface Protein, MSPDBL = Merozoite Surface Protein Duffy Binding Ligand, and GPA = Glycophorin A

2.4.2.2.2 Merozoite Reorientation

The merozoite organelles are concentrated at the apical poles. After merozoite attachment, it reorientates itself bringing its apical ends close to the surface of the erythrocyte (Cowman et al., 2017). This prompts the interaction of the merozoite surface ligands like the EBA and Rh

protein families with the erythrocyte surface receptors resulting in erythrocyte membrane deformation (Weiss et al., 2015). There are five EBA members secreted by the micronemes; the erythrocyte binding antigens 181, 175, and 140 (EBA-181, EBA-175, EBA-140), erythrocyte binding ligand-1 (EBL-1), and the merozoite adhesive erythrocytic binding protein (MAEBL) (Tham et al., 2012). These EBA proteins are essential for erythrocyte invasion, and they facilitate the sialic acid-dependent invasion pathways (Cowman et al., 2017) whereby EBA-175, EBL-1, and EBA-140 bind to sialic acid residues on glycophorin (GP) A, B, and C respectively on the surface of the erythrocyte (Lobo et al., 2003; Maier et al., 2009). Conversely, the merozoite rhoptries secrete members of the Rh protein family such as Rh5, Rh4, Rh2b, Rh2a, and Rh1 which binds to the erythrocyte in a sialic acid-dependent manner (Ord et al., 2015) (Figure 2.6).

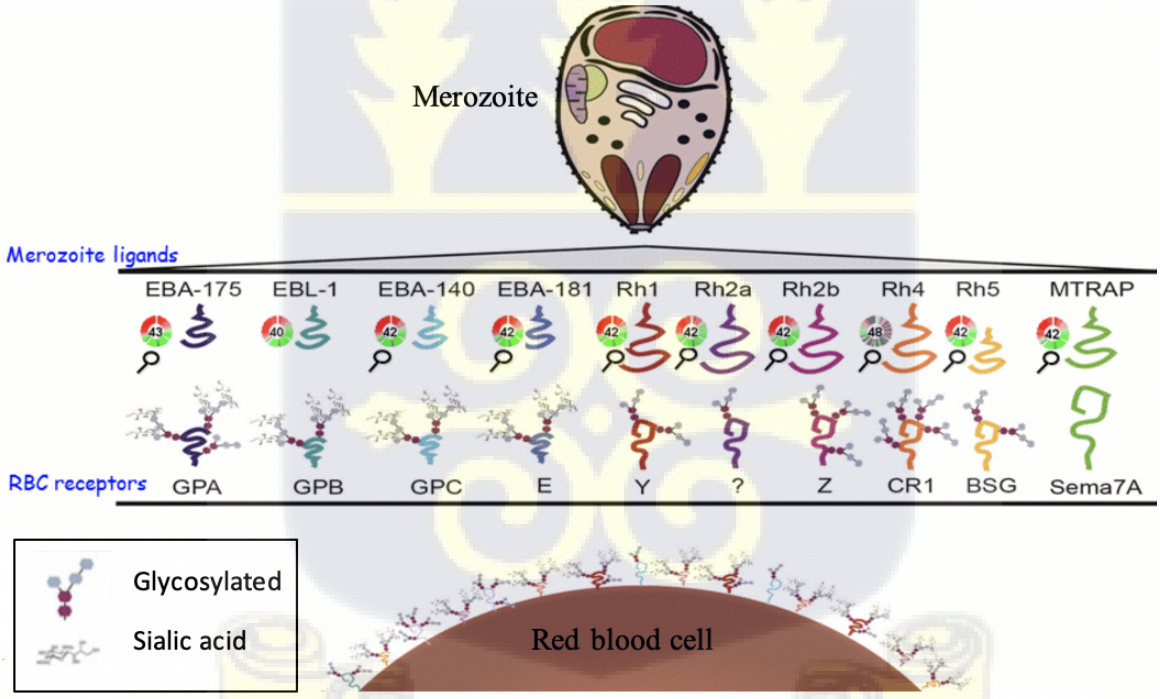


Figure 2.6: Erythrocyte receptors and parasite ligands interactions (Source: Bei & Duraisingh, 2012). EBA = erythrocyte binding antigen, EBL-1 = erythrocyte binding ligand-1, Rh = reticulocyte homolog, CR1 = complement receptor 1, BSG = Basigin, GPA = Glycophorin A, GPB = Glycophorin B, GPC= Glycophorin C, MTRAP = thrombospondin-related anonymous protein

2.4.2.2.3 Glycophorin A (GPA) Receptor and EBA-175 Ligand

A 175 kDa erythrocyte binding antigen, EBA-175, was the first erythrocyte binding protein to be discovered in *P. falciparum* (Camus and Hadley, 1985). EBA-175 is a well-characterized ligand of *P. falciparum* made up of 1,502 amino acids and a highly conserved duplicate (F1 and F2) cysteine-rich region II. EBA-175 has been demonstrated to preferentially bind glycophorin A, a sialoglycoprotein found on the surface of erythrocytes (Sim et al., 1994). The cysteine-rich (Cys-rich) sections F1 and F2 of EBA-175 ligand facilitate merozoite binding to GPA receptor (the major erythrocyte sialoglycoprotein), in a sialic acid-dependent way. EBA-175 cytoplasmic domain (CPD) is also essential in the interaction with GPA and merozoite invasion. The binding of EBA-175 to GPA triggers the release of rhoptry proteins which restores baseline cytoplasmic calcium ($Ca2^{+}$) levels of the merozoite (Singh et al., 2010; Singh & Chitnis, 2017). Also, EBA-175 – GPA interaction causes the cytoskeleton's phosphorylation, which modifies the erythrocyte membrane's characteristics (Sisquella et al., 2017). Moreover, recent knowledge indicates that the interaction between EBA-175 and GPA plays a novel role in supporting erythrocyte clustering by EBA-175 antigen shed post-invasion. This clustering shields merozoites from immune recognition while simultaneously giving them easy access to uninfected erythrocytes (Paing et al., 2018). These changes on the erythrocyte surface result in the transformation of the initial weak and reversible parasite attachment into an irreversible connection known as "tight junction" which facilitates the parasite's total dedication to invading the erythrocyte.

2.4.2.2.4 Glycophorin B (GPB) Receptor and EBL-1 Ligand

The EBL-1 ligand was discovered after EBA-175 (Peterson et al., 1995). It is mostly expressed by mature schizonts and is associated with the fast proliferating phenotype. The single-copy gene EBL-I has two cys-rich domains, a C-Cys domain, and one Duffy-binding-like (DBL) domain with subdomains F1 and F2 in the N-terminal region (Adams et al., 2001; Peterson &

Wellems, 2000). Unlike other EBA members which have eight preserved cysteine residues, EBL-1 only has four. The F2 region of the EBL-1 DBL domain facilitates the binding of erythrocytes during *P. falciparum* invasion (Smith et al., 2000). The human glycophorin B (GPB) is the preferred receptor used by EBL-1 during invasion to bind to the erythrocytes. Thus, the EBL-1 F2 area is a sialic-acid glycan epitope since it interacts with erythrocytes through the GPB receptor, which contains sialic acid residues on its surface (Li et al., 2012; Mayer et al., 2009). Recently, ex vivo experiments have been used to demonstrate that most field isolates and laboratory strains of *P. falciparum* require GPB for invasion (Dankwa et al., 2017).

2.4.2.2.5 EBA-140 and Glycophorin C (GPC)

EBA-140, alternatively called BASEBL is a paralogue of EBA-175. The molecular interactions between the EBA-140 ligand and the erythrocyte receptor are fully described by the structure of this EBA protein family member. Like EBA-175, the cys-rich region II (RII) which consists of the F1 and F2 domains is the receptor binding region (Adams et al., 2001) (Figure 2.7). This EBA-140 binding region is monomeric and unlike other EBA protein family members, the F1 and F2 domains are made up of one glycan binding pocket each of which binds to erythrocyte receptors in different ways, indicating that each domain has a different purpose. Although both domains are necessary for efficient glycan binding, it has been suggested that the F1 domain is the only one that mediates the interaction (Malpede et al., 2013).

EBA-140 binds to GPC (Thompson et al., 2001), a protein expressed on the erythrocyte surface which plays a crucial role in maintaining the shape of the RBC and controlling the components of the cell membrane (Satchwell, 2016). For EBA-140 ligand binding, the GPC region containing amino acid (aa) residues 36 to 63 is essential (Rydzak et al., 2015). Moreover, the extracellular domain of GPC contains one complex N-glycosidic chain at Asn8 and twelve sialylated 0-glycosidic chains attached to threonine and serine residues (Chasis & Mohandas,

1992; Lisowska, 1988). It has been postulated that the N and O-linked oligosaccharides of GPC which are close contain sialic acid residues that group to create a cluster and provide multivalency such that EBA-140 may bind to it effectively.

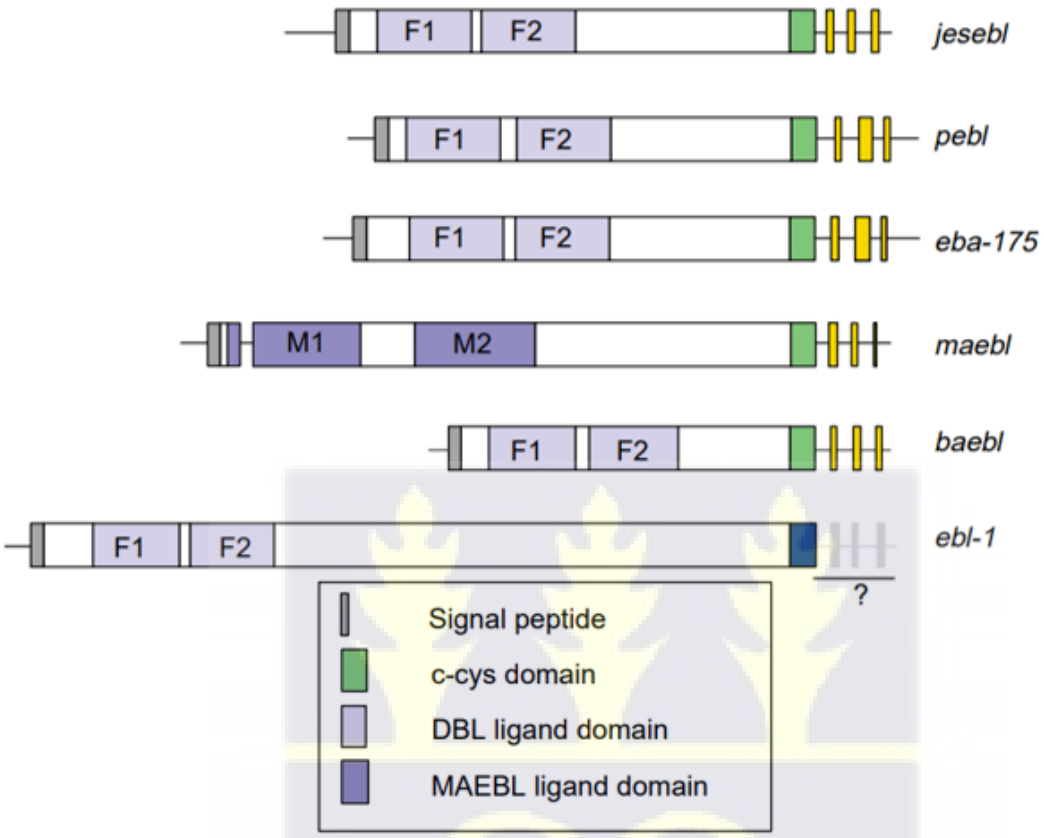


Figure 2.7: Schematic structure of *P. falciparum* erythrocyte binding antigen (EBA) protein family (Source: Adams et al., 2001). c-cys = carboxyl Cys-rich; DBL = Duffy-binding-like, ebl = erythrocyte binding-like, maebl = merozoite adhesive erythrocytic binding protein, jesebl = an alternative name for EBA-181, baebl = an alternative name for EBA-140, pebl = an alternative name for EBA-165.

2.4.2.2.6 Complement Receptor-1 (CR1) and Reticulocyte Homologue-4 (Rh4)

Rh4 is a 220 kDa merozoite ligand (Kaneko et al., 2002) that binds to the human complement receptor-1 during erythrocyte invasion (Tham et al., 2010). CR1 is a glycoprotein made up of a transmembrane region, an N-terminal ectodomain, and a C-terminal cytoplasmic domain,

with several allelic variations. It contributes to the activation of the complement system and the breakdown of immunological complexes. The extracellular domain of the most common allelic variant of CR1 is composed of the complement control protein (CCP) modules. About 28 to 30 CCP modules have been identified, among which only the first three (CCP1-3) bind to Rh4 during the merozoite invasion (Tham et al., 2011).

2.4.2.2.7 Basigin and The Reticulocyte homologue-5 (Rh5)

The merozoite ligand Rh5 is a 63 kDa rhoptry-located protein that interacts with the receptor Basigin (a 43-66kDa immunoglobulin superfamily transmembrane glycoprotein) during erythrocyte invasion (Baum et al., 2009; Crosnier et al., 2011). This interaction causes calcium (Ca^{2+}) influx to the erythrocyte known to facilitate the activation and release of merozoite proteins that are reoriented apically to the erythrocyte (Weiss et al., 2015). Despite lacking a transmembrane domain and a GPI anchor signal, the Rh5 and Basigin interaction is used by all *P. falciparum* strains during erythrocyte invasion (Crosnier et al., 2011; Douglas et al., 2011). Hence, erythrocyte invasion is facilitated by the formation of a complex between reticulocyte binding protein homologue 5 interacting protein (Ripr), cysteine-rich protector antigen (CyRPA), and P113 (Reddy et al., 2015). CyRPA and Ripr are other invasins secreted from the micronemes. The complex is formed when micronemal protein CyRPA binds to the N-terminal of Rh5 and another micronemal protein Ripr in turn binds to CyRPA (Chen et al., 2017; Wong et al., 2019). Rh5 also interact with P113 and this interaction has been suggested to enable the anchoring of the complex to the parasitophorous vacuolar membrane during invasion (Galaway et al., 2017). Thus, Rh5-basigin and their tripartite interaction with CyRPA, Ripr, and P113 has emerged as a possible vaccine target (Douglas et al., 2015; Jin et al., 2018).

2.4.2.2.8 Tight Junction Formation

The third phase of invasion is characterised by the creation of a tight junction at the point of contact between the apical tip and the erythrocyte using proteins secreted from the microneme

and rhoptry organelles. The binding of AMA1 to a group of rhoptry neck proteins (RONs) represents an important receptor-ligand interaction at this point (Srinivasan et al., 2011). Ron 2, 4, 5, and 8 are moved into the erythrocyte membrane to respond to upstream stimuli or interact with AMA1. While RON-2 is inserted as a membrane protein and serves as the membrane antigen's anchor, RON-4, 5, and 8 appear to adhere the complex to the cytoskeleton of the erythrocyte (Besteiro et al., 2009). It has been shown that the interaction of AMA1 and RON-2 is essential for *P. falciparum* invasion because it is responsible for the creation of the tight junction (Lamarque et al., 2011; Srinivasan et al., 2011). Hence, it has been shown that targeted blockage of the AMA1-RON-2 complex prevents junction creation. Therefore, the tight junction formation is regarded as the final stage for a coordinated invasion process because it initiates all subsequent, sequential invasion processes, with the contents of rhoptries being secreted in a coordinated manner (Riglar et al., 2011).

2.4.2.2.9 Active Invasion

A single-headed myosin anchored to the erythrocyte membrane is necessary for a successful invasion because it drives short actin filaments in a single direction. The invasins, a member of the thrombospondin-related anonymous protein (TRAP) family, are connected to the actin filaments. The TRAP proteins engage with the erythrocyte receptors extracellularly, causing the parasite to progress by turning on the actin-myosin motor (Bartholdson et al., 2012). The micronemal protein MTRAP (merozoite-specific TRAP), a conserved merozoite TRAP homologue has been discovered to bind the actin-binding protein aldolase, making it possible for the merozoite to invade using actin polymerisation and a homologous motor (glideosome) (Baum et al., 2006; Morahan et al., 2009). The ectodomain of MTRAP interacts with Semaphorin -7A (a receptor for MTRAP) (Bartholdson et al., 2012). In turn, this MTRAP-Sema-7A interacts with AMA1-RONs to form a stable, high avidity ring-like structure that functions as an anchor and supplies the necessary pulling force for the merozoite to penetrate

the membrane of the parasitophorous vacuole (Yap et al., 2014). The parasite uses the vacuole to isolate itself from the host cell's cytoplasm and create a favourable environment for its growth (Weiss et al., 2016).

During the first 18 hours in the erythrocyte, the parasite maintains its ring-like shape and does not change much morphologically. It subsequently develops into a feeding stage (trophozoite), where it grows quickly and becomes more metabolically active. At the schizont stage, DNA synthesis and nuclear division produce 16–32 merozoites (Cowman et al., 2017). Rapid cytokinesis then occurs, allowing distinct organelles and nuclei required for invasion to produce distinct merozoites which are released (through a process known as egress) to infiltrate fresh erythrocyte for the cycle to continue (Weiss et al., 2015).

2.4.2.3 Experimental Determination of *P. falciparum* Erythrocyte Invasion Pathways

The interaction between the EBA and Rh proteins and erythrocyte surface receptors are known as invasion pathways. Although extensive research has been done on these protein families and some corresponding receptors have been identified, others have remained mysterious. Several approaches have been used to study *P. falciparum* invasion pathways such as the use of antibodies against merozoite ligands or antibodies against erythrocyte receptors (Gaur et al., 2007). In investigations to comprehend erythrocyte invasion routes, erythrocyte binding assays have been utilised to demonstrate the binding of parasite ligands to erythrocytes (Gaur et al., 2007; Mayer et al., 2002). Currently, the most common experimental approach to studying parasite invasion pathways is using enzymes like neuraminidase, trypsin, and chymotrypsin to artificially deplete erythrocyte receptors. Neuraminidase cleaves sialic acid residues found on the glycophorin (GP) receptors. Since sialic acid residues are required for merozoite binding, cleavage can prevent erythrocyte invasion. However, several surface proteins have their peptide backbone broken by endopeptidases like trypsin and chymotrypsin (Baum et al., 2003). Trypsin treatment of erythrocytes breaks down GPA, GPC, and CR1, whereas chymotrypsin

treatment eliminates band 3, GPB, and other proteins linked to invasion (Cowman & Crabb, 2006; Triglia et al., 2011)

Six well-characterized *P. falciparum* invasion phenotypes have been defined according to their level of sensitivity to the three commonly used enzymes (Cowman et al., 2017). Isolates with invasion inhibition greater than 50 % after enzyme treatment compared to no enzyme-treated erythrocyte control (RPMI) are considered sensitive (S) while those with invasion inhibition below 50 % are resistant (R) (Baum, Pinder, et al., 2003; Lobo et al., 2004). Hence, two types of erythrocyte invasion routes exist those that rely on sialic acid (SA) (SA-dependent) and those that do not (SA-independent). Based on combinations of sensitivity or resistance to various enzymes, the following pathways are feasible: NsTsCr (EBA175 - GPA and EBA140-GPC receptor), NsTrCs (EBL1 - GPB and EBA181 - E receptor), NsTrCr (Rh1 - Y receptor), NrTsCs (Rh4 - CR1 receptor), NrTrCs (Rh2b - Z receptor), NrTsCr (uncharacterized), NsTsCs (uncharacterized), and NrTrCr (uncharacterized) (Table 2.1), where N, T, and C stands for neuraminidase, trypsin and chymotrypsin respectively. Of these, NsTsCr and NsTrCr are SA-dependent while the rest are alternative SA-independent patterns.

Table 2.1: Characterised invasion pathways of *P. falciparum* strains and their enzyme sensitivity

Parasite ligand	**EBA 175**	**EBL-1**	**EBA 140**	**EBA 181**	***Pf*Rh 1**	***Pf*Rh 2**	***Pf*Rh 4**	***Pf*Rh5**
Erythrocyte receptor	**GPA**	**GPB**	**GPC**	**E**	**Y**	**Z**	**CR1**	**Basigin**
Chymotrypsin	R	S	R	S	R	S	S	R
Trypsin	S	R	S	R	R	R	S	S
Neuraminidase	S	S	S	S	S	R	R	R

S = sensitivity, R = resistant, GP = glycophorin, EBA = erythrocyte binding antigen, EBL = erythrocyte binding ligand, and CR1 = complement receptor (Duraisingh et al., 2003).

Advances in microscopy and flow cytometry technologies have resulted in the development and use of sensitive high throughput assays to charaterise these pathways. Invasion phenotypes are commonly determined with the use of a two-coloured flow cytometry method whereby, equal amounts (1:1 ratio) of unlabelled infected donor erythrocytes are cultured with a targeted population of labeled uninfected erythrocytes pre-treated with either neuraminidase, trypsin and/or chymotrypsin. This enables the determination of invasion phenotypes through the quantification of heterologous reinvasion (parasite reinvasion into the target erythrocytes) over a single erythrocytic cycle (Bei & Duraisingh, 2015).

2.4.2.4 Alternate Erythrocyte Invasion Pathways

It is well established that *P. falciparum* laboratory strains have varied invasion pathways. For instance, HB3, 7G8, and 3D7 strains use the SA-independent invasion pathways (Baldwin et al., 2014) while w2mef and Dd2 use the SA-dependent pathways (Dolan et al., 1990). Studies on clinical isolates in Sub-Saharan Africa have demonstrated the use of multiple invasion pathways and shown that the invasion pathways diversity in clinical samples is similar to that observed in laboratory strains (Deans et al., 2007; Gomez-Escobar et al., 2010; Mensah-Brown et al., 2015). Hence the parasites are capable of switching from one invasion pathway to another (Awandare et al., 2018). Although the dynamics of erythrocyte invasion and the selection for alternative invasion phenotypes are not fully understood to date, it is believed that these processes are influenced by factors like transmission intensity and allelic diversity, erythrocyte heterogeneity, nutritional, and immunological responses, drug pressures and pressure from different control interventions that affect gene expression or result in genetic polymorphisms (Josling et al., 2015). Therefore continuous surveillence is needed to know which pathways the parasites are using for invasion at a particular time and what genes are expressed so as to better determine what control interventions are needed to achieve the zero malaria target.

2.4.2.5 Variation in Invasion Pathways due to Differential Gene Expression

Like most eukaryotic cells, the transcription machinery of *P. falciparum* contributes to controlling the parasite's gene expression. It is itself modulated by environmental stressors like temperature changes and changes in glucose levels (Hughes et al., 2010). Transcription at the intraerythrocytic cycle of the malaria parasite is a well-organised process, tied to the parasite's demands during each step of the cycle (Bozdech et al., 2003). The majority of genes are expressed during the trophozoite stage, and these genes are involved in protein transport, protein catabolism, and metabolism, among other parasite functions. However, genes that code for proteins involved in the biosynthesis of micronemes and rhoptries are expressed during the schizont stage of the erythrocytic cycle (Lu et al., 2017). Thus, the invasion phenotype of each parasite line depends on a molecular hierarchy that governs which of the transcribed ligands are employed (Baum et al., 2005).

Experimentally, it has been shown that variant expression of invasion ligand genes can result in changes in *P. falciparum* erythrocyte invasion pathways (Duraisingh, Maier, et al., 2003; Triglia et al., 2004) depending on the environment of the specific host (Stubbs et al., 2005) or whether the culture is static or suspended (Awandare et al., 2018). Variable expression has been observed in all members of the Rh and EBA families. Since each member of these two protein families interacts with different receptors on the surface of the erythrocyte, altering their expression might give the parasite the flexibility to invade various types of erythrocytes including mutant erythrocytes with altered surface protein composition and those at various stages of maturity (Cortés, 2008). Clonally variable gene expression in malaria parasites is one source of transcriptional variability. Given that spontaneous variations in expression result in transcriptionally heterogeneous populations, clonally variable gene expression offers a clear potential for adaptation. In population biology, it is well known that heterogeneous populations are more adaptable than homogeneous populations in environments that are constantly

changing (Kussell & Leibler, 2005). This is because diversity creates the conditions for natural selection of individuals with phenotypes that confer the greatest fitness upon an environmental change. The first description of clonally variable gene expression in *P. falciparum* was the var genes (Scherf et al., 2008), a group of roughly 60 genes that encodes PfEMP-1, a crucial component of pathogenesis and immune evasion. Following that, the expression of other gene families, like those involved in erythrocyte invasion (e.g DBLMSP2) have also been shown to be clonally variable (Amambua-Ngwa et al., 2012). For instance, studies have shown that the genetic disruption of Rh1, EBA-175, and Rh2b resulted in some parasite lines using alternative invasion pathways than their parental lines (Duraisingh, et al., 2003; Triglia et al., 2004). These necessitate continuous research on the transcriptional profile of invasion ligands especially in very low edemic countries like The Gambia in order to determine what genes are constantly express and how this affect malaria control and elimination.

2.4.2.6 Variation in Invasion Pathways Due to Genes Polymorphisms

Genetic polymorphisms in invasion ligands have been suggested to contribute to variation in erythrocyte invasion pathways and serve as a means of immune evasion by the parasite. Likewise, all merozoite ligands that are differentially expressed are polymorphic and some of their sequences exhibit the hallmarks of diversifying selection (Iriko et al., 2008; Verra et al., 2006). For instance, single nucleotide polymorphisms (SNPs) in EBA-175's Region II have been shown to influence its binding to several receptors, pointing to a possible mechanism involving genomic sequence polymorphism that accounts for variability in invasion pathways (Mayer et al., 2002). In vitro, polymorphism of EBA-181 and EBA-140 can change their receptor specificity (Mayer et al., 2002, 2004). Moreover, it has been demonstrated that Rh5 polymorphism alters host receptor recognition, which has an impact on the capacity of human *P. falciparum* strains to infect *A. nancymaae* monkeys (Hayton et al., 2008). Hence, genetic studies aimed at identifying these polymorphisms in both the parasite ligands and the host

receptors are necessary to understand the parasite's biology, genetic diversity, host-parasite interactions as well as identify novel potential targets for drugs and vaccines development.

2.4.3 Parasite Sequestration

Sequestration is a process whereby a red blood cell infected with mature *P. falciparum* stages (trophozoites and schizonts) is attached to endothelial cells, allowing only the ring stages to circulate in the blood (White et al., 2013). This process described over a century ago is used as a survival mechanism by the parasite because it helps the parasite to avoid splenic clearance (Rowe et al., 2009). Sequestration may have disastrous effects on the host. For instance, cerebral malaria, the primary cause of death from *Plasmodium* infections, has a fundamental pathophysiological hallmark of extensive sequestration in the microvasculature of the brain (Storm et al., 2019). Moreover, sequestration in the placenta by parasites that expresses the *P. falciparum* erythrocyte membrane protein-1-variant surface antigen-2 chondroitin sulphate A (PfEMP1-VAR2CSA) results in significant clinical effects on both the mother and the unborn child (Resende et al., 2008). Cytoadherence, rosetting, and auto agglutination are some events used to explain the process of sequestration in *P. falciparum*.

2.4.4 Cytoadherence

Infected erythrocytes adhere to microvascular endothelial cells in a variety of organs, including the placenta, intestine, liver, and brain through a process called cytoadherence. Three different groups of parasite-derived variant surface antigens (VSAs) encoded by multigene families, including PfEMP1 (Jensen et al., 2004), repetitive interspersed family (RIFIN) proteins (Goel et al., 2015), and subtelomeric variable open reading frame (STEVOR) proteins (Niang et al., 2014) are used to cytoadhere infected RBCs. The endothelial cell ligands known to interact with these *P. falciparum* surface proteins include the intercellular adhesion molecule-1 (ICAM1) (Smith et al., 2000), CR1, platelet/endothelial cell adhesion molecule (PECAM),

endothelial protein C receptor (EPCR), chondroitin sulphate A (CSA) and heparin sulphate (Lau et al., 2015; Turner et al., 2013) (Figure 2.8).

Apart from its role in immune evasion, cytoadherence also contributes to the aetiology of severe malaria. For instance, the interaction of PfEMP1 with EPCR is associated with cerebral malaria aetiology, one of the most significant types of malaria-related sequelae (Kessler et al., 2017; Mkumbaye et al., 2017).

2.4.5 Rosetting

Rosetting is the binding of an infected erythrocyte to two or more uninfected erythrocytes to form a group of cells that resembles a flower, resulting in micro-vascular obstruction (Vogt et al., 2006). Just like cytoadherence, studies on human postmortem have demonstrated that rosetting is frequently linked to severe pathogenesis, including severe malaria anaemia and cerebral malaria (Doumbo et al., 2009; Rowe et al., 1995). Only parasites capable of forming rosetting are linked to severe pathogenesis and the degree of rosette formation in *P. falciparum* field isolates differ from laboratory strains. Moreover, rosette formation also helps the parasite to hide its VSAs expressed on the infected erythrocyte surface with uninfected red blood cells enabling them to avoid immunological identification by antibodies or phagocytes (Lee et al., 2019).

STEVOR, RIFIN, and PfEMP1 are known to be implicated in the mechanisms of *P. falciparum* rosetting. These proteins interact with different host-derived receptors; most of which interact with the different extracellular domains of PfEMP1 (Figure 2.8). The presence of PfEMP1 rosetting ligands (CR1 and herapan sulphate) on endothelial cells suggests that the main function of PfEMP1 is possibly endothelial adhesion, which prevents clearance of infected erythrocytes in the spleen. This implies that (Figure 2.8i) infected erythroctes use PfEMP1 to adhere to the endothelial receptors, (Figure 2.8ii) uninfected RBCs use the RIFIN or STEVOR to bind to cytoadhered iRBC (Figure 2.8iii) rosetting then contributes to

sequestration by attaching to cytoadhered non rosetting parasites with PfEMP1, RIFIN, and STEVOR causing vascular blockage and escalating pathology and disease severity. In every scenario, the rosette is anticipated to be the parasite against the host's immunological reaction and to offer a conducive environment for newly released merozoites following schizont rupture to facilitate effective invasion (Figure 2.8iv) (Goel et al., 2015; Yam et al., 2017).

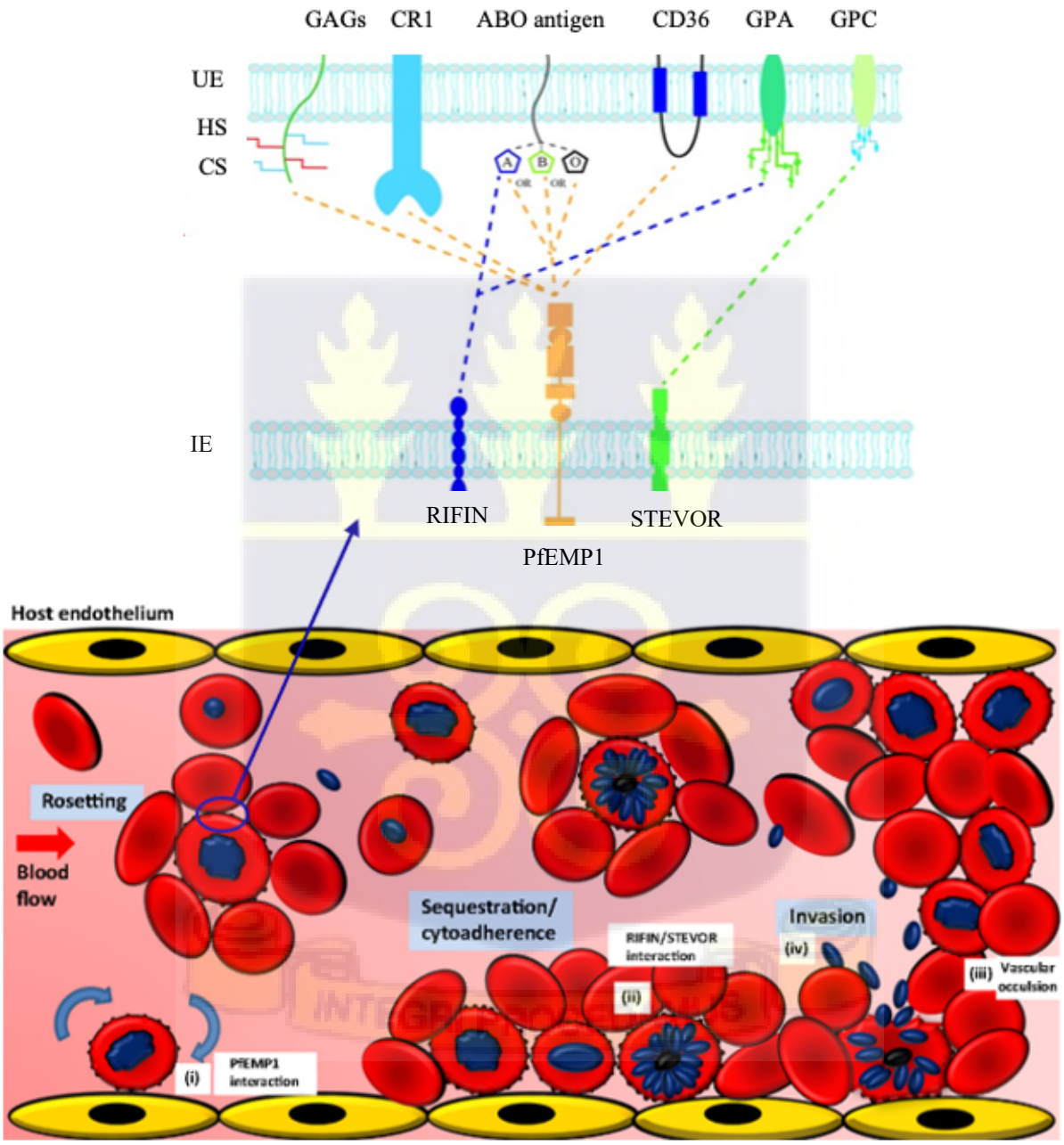


Figure 2.8: Sequestration, cytoadherence, and resetting of *P. falciparum* (Adapted from McQuard and Rowe, 2019; Yam et al., 2017). GPA = glycophorin A, GPC = glycophorin C,

CR1 = complement receptor 1, HS = heparin sulphate, UE = uninfected erythrocyte, IE = infected erythrocyte, GAGs = glycosaminoglycans, CS = chondroitin sulphate, CD = Cluster of differentiation, RIFIN = PfEMP1 = *P. falciparum* erythrocyte membrane protein-1, STEVOR = Subtelomeric Variable Open Reading frame. Dotted lines indicate suggested receptors for each ligand.

2.4.6 Erythrocyte Disorders and their Association with Malaria

Throughout human history, several genetic mutations have emerged and have been selected in world populations due to diseases or other biological and environmental conditions. While the most deleterious mutations often vanish from the population, some confer a survival benefit to the carrier and have increased in frequency over multiple generations due to positive selection. Some examples of these positively selected mutations are blood disorders associated with malaria pathogenesis like thalassaemias (α and β thalassaemia), sickle cell disease (HbC, HbE, and HbS), the Duffy-negative blood group, and glucose-6-phosphate dehydrogenase (G6PD) deficiency (Kariuki & Williams, 2020) (Figure 2.9). Additionally, advances in research approaches and the completion of the human genome project have led to the discovery of more loci on the human genome linked to malaria severity. Some of these recently described markers include structurer variations at the glycophorins which act as receptors for malaria parasite invasion (Leffler et al., 2017). Other genetic factors like pyruvate kinase, ovalocytosis, nitric oxide synthase 2, haem oxygenase, haptoglobin, ATP2B, and ABO blood groups have also been associated with malaria susceptibility either by altering immune responses or by disrupting host-parasite interactions. These have shed light on the dual selection and co-adaptation of polymorphisms that occur in the parasite and its human host (Driss et al., 2011).

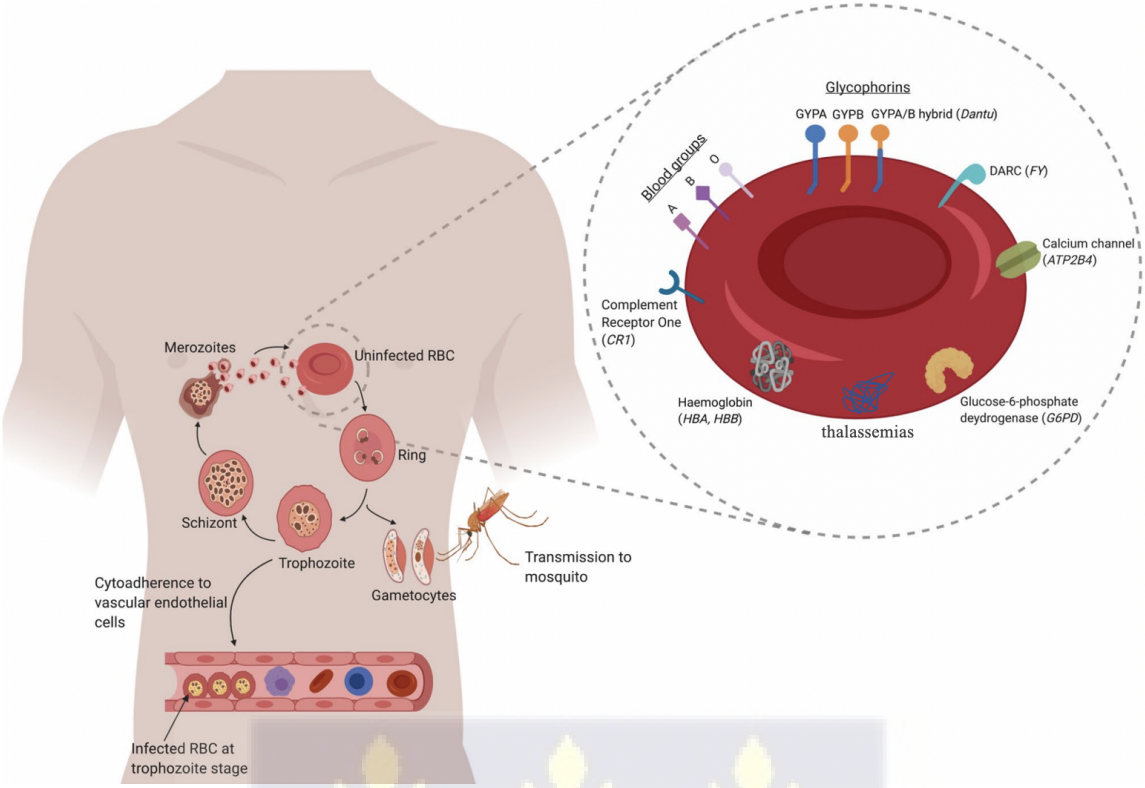


Figure 2.9: The erythrocytic cycle of *P. falciparum* indicating red blood cell (RBC) variants associated with malaria (Source: Kariuki & Williams, 2020). GPA = glycophorin A, GPB = glycophorin B, CR1 = Complement receptor 1, DARC = Duffy Antigen/Receptor for Chemokines, HBA = Hemoglobin A.

2.4.7 Haemoglobinopathy

The thalassemias are caused by mutations in either the α-globin or β-globin genes. There are four α-globin genes in humans, two on each chromosome. Therefore, mutations can result in an individual having either three or fewer α-globin genes, leading to anaemia. This condition is termed α-thalassemia and the fewer the α-globin genes, the more severe the anaemia gets but this is not life-threatening. The thalassemias (alpha and beta thalassemia) are the most prevalent monogenic blood disorders in Africa due to their protective effect against malaria (Farashi & Harteveld, 2018; Weatherall & Clegg, 2001). Although several studies have reported and proven that α-thalassemia protects against malaria, there are some contradictions in these

reports especially regarding the malaria phenotype. There are reports that it protects against mild and severe malaria (Mockenhaupt et al., 2004; Williams et al., 2002), and it has been shown that it is protective against mild malaria through decreased rosetting (Williams et al., 1996). Homozygotes with higher micro-erythrocyte counts have less hemoglobin loss for a given parasite density and thus prevent the development of severe anaemia (Fowkes et al., 2008).

2.4.8 Glucose-6-phosphate dehydrogenase (G6PD) deficiency

G6PD deficiency is the lone source of reduced nicotinamide adenine dinucleotide phosphate (NADPH) in the erythrocyte (Luzzatto & Arese, 2018). NADPH is essential for glutathione (an antioxidant) homeostasis. Thus, G6PD deficiency indicates an unstable amount of NADPH which in turn affects the ability of the RBCs to maintain glutathione homeostasis, resulting in oxidative stress and RBC hemolysis. G6PD deficiency is a genetic X-linked haemolytic condition with incomplete inheritance. It is the most common enzymopathy in humans, highly present in the tropics and subtropics, synchronising with malaria endemicity or previously endemic areas in Africa, Asia, Latin America, and the Mediterranean (Howes et al., 2012). This implies that the distribution of G6PD deficiency is through natural selection by malaria (Kwiatkowski, 2005). G6PD deficiency protects against malaria through the process of phagocytosis whereby, ring-parasitised variant RBCs are phagocytosed more often because of increased oxidative stress. Most *P. falciparum* malaria studies in Africa have supported this protective effect of G6PD deficiency against complicated malaria (Clarke et al., 2017). However, there have also been conflicting views on this topic. According to some studies, G6PD deficiency have no protective effect against malaria, while others suggest that protection is observed in only female heterozygotes, or male hemizygotes (Clarke et al., 2017; Mbanefo et al., 2017; Uyoga et al., 2015).

2.4.9 Haemoglobin S

The main protein of the RBC and the oxygen-carrying component, haemoglobin, is typically synthesized as a tetramer of two alpha and two β-globins, known as adult haemoglobin A (HbA). Sickle cell disorders (SCD) are a class of inherited illnesses caused by mutations in the globin gene and are often defined by painful Vaso-occlusive crises and chronic haemolytic anaemia (Rees et al., 2010). It results from a point mutation in one of the two chains of β-globin whereby Glutamate (Glu) is substituted with valine (Val) at position six, resulting in an altered protein called haemoglobin S (HbS) (Ingram, 1958). The epidemiology of SCD is parallel to that of malaria particularly in Sub-Saharan Africa with high malaria burden (Piel et al., 2013; Wiebe et al., 2017). This is not a coincidence because studies have shown that people with the heterozygous form of the sickle (Hb genotype AS; HbAS) have some level of protection against severe malaria while homozygous (HbSS) individuals are at higher risk of mortality due to malaria (Luzzatto, 2012; Makani et al., 2010). Several mechanisms by which HbAS offers protection against malaria have been described. Some of these include increased splenic clearance of sickled RBCs, enhanced phagocytosis of ring-parasitised variant erythrocytes, acquired host immunity, decreased cytoadherence, resetting, decreased trafficking of parasite proteins to the surface of RBCs, and suppression of parasite development caused by oxygen-dependent polymerization of sickle haemoglobin (Archer et al., 2018; Williams et al., 2005).

2.4.10 Haemoglobin C and E (HbE)

Haemoglobin C (HbC) occurs because of a mutation in the amino acid sequence of the beta-globin gene at position six where glutamic acid is replaced by lysine (β6Glu - Lys). Individuals with the heterozygous form (HbAC) are asymptomatic while those with the homozygous form (HbCC) often suffer from mild symptoms like splenomegaly, haemolysis, or gallstones. Just like HbS, HbC has also been associated with malaria phenotypes. It has been suggested that HbC provides substantial protection against severe malaria while HbAC offers minimal

protection (Taylor et al., 2012). A study in Burkina Faso suggested that children with the haemoglobin C trait are protected from both complicated and uncomplicated malaria (Modiano et al., 2001). This finding was however contradicted by a similar study in Ghana which observed no protection against both mild and severe malaria in individuals with haemoglobin C trait (Kreuels et al., 2010). Thus, due to these contradictory findings, it is unknown whether HbC gives relative resistance to disease, infection, or both. Although the exact mechanism of protection is not clear, various mechanisms have been suggested such as (1) aberrant appearance of *P. falciparum* erythrocyte membrane protein 1 (PFEMP-1) which reduces cytoadherence and possibly reduces the ability of the parasite to sequester and avoid immune clearance (Fairhurst et al., 2005), (2) rapid development of malaria immunity (Verra et al., 2007), and (3) abnormal *P. falciparum* intraerythrocytic development resulting to lower replication rates in some HbC individuals (Fairhurst et al., 2003).

Another haemoglobin variant, haemoglobin E (HbE) has also been associated with malaria severity. HbE variant is caused by a point mutation in the amino acid sequence of the beta-globin gene at codon 26 where glutamic acid is replaced by lysine (Fucharoen & Weatherall, 2012). Heterozygous individuals (HbAE) have moderate anaemia, while homozygotes (HbEE) suffer from microcytic hypochromic anaemia (Kohne, 2011). Some suggested mechanisms of HbE protection against malaria severity include the decreased likelihood of merozoite invasion into RBCs (Chotivanich et al., 2002), inhibiting the remodeling of host cell's structure and parasite's development after a successful erythrocytic invasion (Chotivanich et al., 2002; Lelliott et al., 2015), enhanced phagocytosis by monocytes (Bunyaratvej et al., 1986), and high antimalarial antibody titres (O'Donnell et al., 2009).

2.5 Erythrocyte Receptor Polymorphisms and Association with Malaria

2.5.1 Overview of the Human Glycophorins

The glycophorins are low molecular weights transmembrane type 1 glycoproteins of the erythrocytes containing sialylated N-glycans and/or O-glycans (Chasis & Mohandas, 1992; Lisowska, 1988b). Glycophorins A, B, C, and D are the four main glycophorin proteins found in humans. Two similar genes GPA and GPB code for glycophorins A and B, while an unrelated gene GPC codes for glycophorins C and D. Separate translational start spots on the GPC transcript account for the differences between glycophorin C and D (Kim et al., 1987). Additionally, glycophorin C is considered a human-specific protein because it has a translation initiation codon that is unique to humans whereas glycophorin D is present among all apes due to its conserved translation initiation codon across apes (Wilder et al., 2009). There is another annotated glycophorin gene like GPA and GPB with no detected protein, termed GPE (Vignal et al., 1990). GPE has been suggested to be a pseudogene due to a lack of expression on the surface of the erythrocyte. However, a recent study in Japan reported for the first time, the expression of a GPE (Tsuneyama et al., 2020), indicating that GPE might not be a pseudogene as initially suggested. Three of the glycophorin genes (GPA, GPB, and GPE) constitute a 350-kb gene cluster on chromosome four, while GPC is found on chromosome two (Lopez et al., 2021; Mattei et al., 1986). The glycophorins are known to carry different blood groups. For instance, the Ss and MN blood groups are determined by GPA and GPB whereas the Gerbich blood group system associated with malaria severity is determined by GPC (Lopez et al., 2021). Given the large number of glycophorins on erythrocytes surfaces, it is possible that these molecules also act as a substrate for glycosylation, resulting in a negatively charged, complex glycan coat that hinders the RBCs from sticking to other cells during circulation.

2.5.2 Glycophorin Genes Structural Variation

Three 120 kb tandemly organized repeats on chromosome four of the human genome make up the glycophorin gene locus (Figure 2.10A). Each repeat carries one of three closely related (about 97% sequence homology) glycophorin genes: glycophorin E (GPE), glycophorin B (GPB), and glycophorin A (GPA) (Onda et al., 1994; Vignal et al., 1990) which are vulnerable to rearrangements due to frequent non-allelic homologous recombination (NAHR) events. NAHR events can either be duplications, insertions, deletions, or fusions (Figure 2.10B). Research has shown that GPB and GPE originated from GPA (the ancestral gene) due to a sequence of molecular events (Onda et al., 1993) beginning with the duplication of GPA. This was followed by misalignment of two chromatid strands each containing a replicated GPA. Then, unequal crossing-over took place within the Alu sequences found in each strand, resulting in the generation of a progenitor GPB/GPE genomic region (Onda et al., 1993). The progenitor GPB/GPE genomic region was later duplicated, resulting in the independent emergence of the GPB and GPE genes. Unlike GPA, B, and E, GPC is not vulnerable to rearrangements or structural variation because it is a single-copy area on chromosome two. However, there have been reports of deletions of either exon two, exon three, or both resulting in the Gerbich-negative blood types Ge2, Ge3, and Ge4 respectively. Additionally, the duplication or triplication of GPC exon three results in Gerbich Lsa antigen (Figure 2.10B) (Jaskiewicz et al., 2018).

Recent studies have used genomic sequence data to systematically catalog structural variants in the glycophorin gene locus across Sub-Saharan Africa. Several deletions (DEL) and duplications (DUP) have been identified in this region. Either of the glycophorin genes can be deleted either partially or completely to give rise to a new or null phenotype. Deletion of GPA results in a phenotype called En (a-), while deletion of GPB results in the U- phenotype, and deletion of both GPA and GPB gives rise to the M^kM^k phenotype. GPB deletions are the most

common variants and based on their prevalence, they have been designated numbers from one to eight with GPB deletion 1 (DEL 1) being the most prevalent and GPB DEL 8 being the least. Among these, deletions 1 and 2 are more prevalent in West Africa. Same as the deletions, duplications have also been numbered from 1 to 8 based on their prevalence, among which duplication 4 (DUP4) found only in East Africa is the most characterised and forms the basis of the Dantu NE blood group (Leffler et al., 2017; MalariaGEN, 2015). Molecular approaches like PCR, and fibre-FISH (flourescence in situ hybridization) among others have been used to validate some of the variants and some have been identified to be the driving force of blood grouping variation in humans. However, most of the identified variants in the human glycophorin locus are yet to be characterised and their association with malaria established. This necessitate reaserach in this area with the aim to genotype and characterised some of these variants especially the most common once like the glycophorin B deletions 1 and 2.

Three molecular mechanisms that account for variant glycophorin diversity includes gene conversion, unequal crossing-over, and single nucleotide polymorphism (Blumenfeld & Huang, 1997).

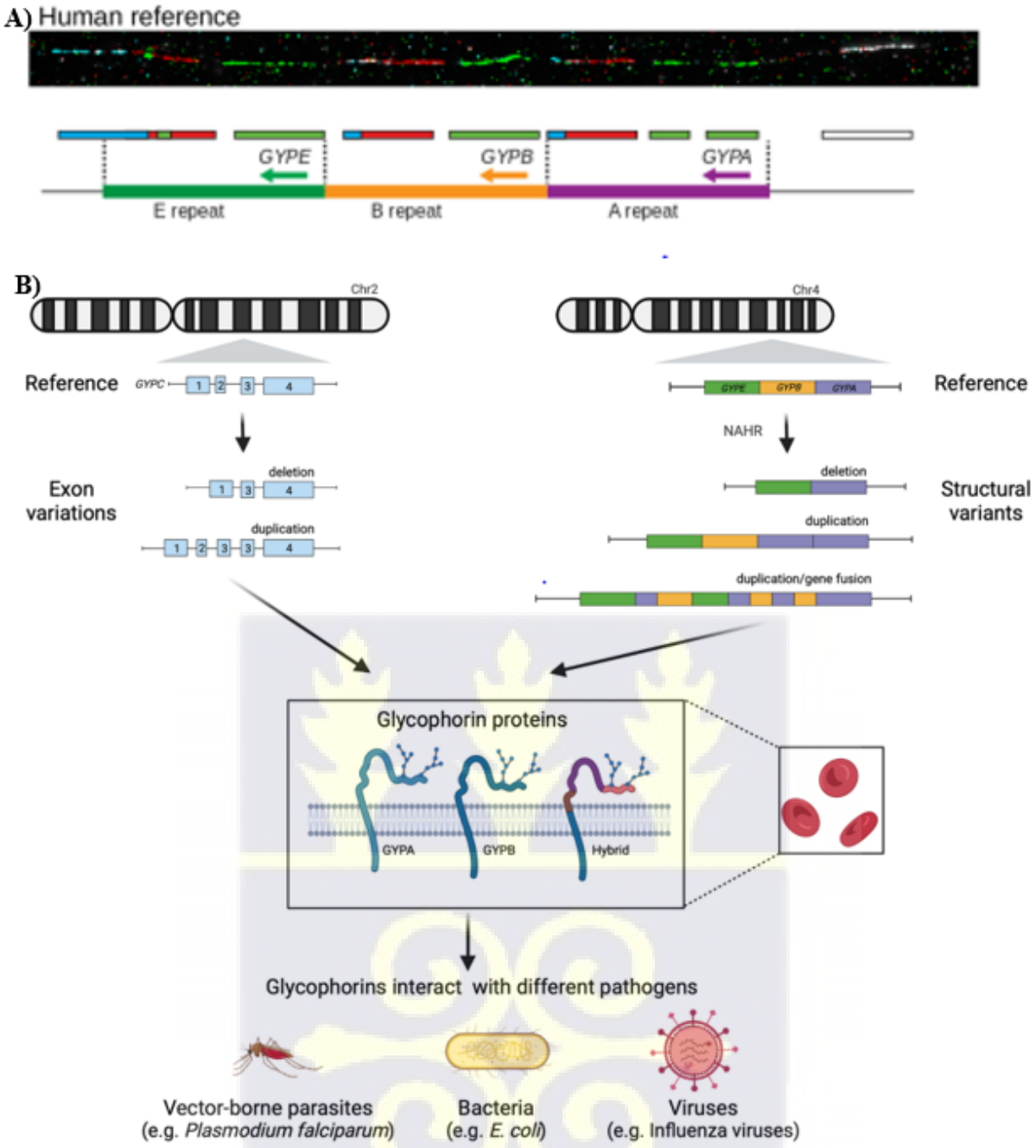


Figure 2.10: The glycophorins and their role in vector-borne parasites (Source: Hollox & Louzada, 2023). A) Typical representation of the reference haplotype in the human glycophorin locus. Green-, orange-, and purple-coloured bars represent 120-kb repeats conveying GPE (glycophorin E), GPB (glycophorin B), and GPA (glycophorin A), respectively. B) GPC (glycophorin C) and the GPA/GPB/GPE genes are located on two

separate loci. On the surface of the erythrocyte, many variants encode different glycophorin versions which interact differently with pathogens like malaria, viruses, and bacteria.

2.5.2.1 Gene Conversion

The process of gene conversion occurs during meiosis due to the replacement of homologous sequences in the acceptor chromatid strand with nucleotide sequences from a donor chromatid strand (Kobayashi, 1992; Lorenz & Mpaulo, 2022). It occurs through a double-strand break repair mechanism whereby a short fragment of a gene is replicated into another gene during gene duplication (Kobayashi, 1992). There is no reciprocity in this genetic material transfer. GP (A-B-A) and GP (B-A-B) hybrid genes are created via gene conversion between GPA and GPB (Palacajornsuk, 2006; Reid, 2009). In GP (B-A-B) hybrids, this repair mechanism modifies a section in GPB corresponding to GPA exon three so that an inactive GPB splice site sequence is replaced by an active sequence of GPA, resulting in the creation of a hybrid protein from the translation of a partial GPB called GPB pseudoexon three. Additionally, GPE recombines with GPA and GPB to create GP (A-E-A) and GP (B-E-B) respectively (Tsuneyama et al., 2020; Willemetz et al., 2015).

2.5.2.2 Unequal crossing-over

GP (B-A) and GP (A-B) hybrid alleles are created through unequal crossing-over. Unequal crossing over occurs because of the misalignment of glycophorin genes during meiosis. During this process, two hybrid genes are produced in reciprocal configurations by the exchange of genetic material of uneven length between the two sister chromatid strands. As a result, one strand (Lepore type) acquired less genetic material while the other strand (anti-Lepore type) received more genetic material than each originally provided (Palacajornsuk, 2006).

2.5.3 The Human Glycophorin Variants and Malaria

Malaria creates tremendous selective pressure on the human genome and studies on genetic variations have identified signatures of natural selection in the GPA-GPB-GPE region (Bigham

et al., 2018; Johnson & Voight, 2018; Ko et al., 2011). Several structural variants like duplications, deletions, insertions, and fusion have been identified in the glycophorin locus, and based on their prevalence in the population, some deletions (DEL) and duplications (DUP) have been numbered from 1 to 8 where 1 is the most prevalent and 8 is the least (Leffler et al., 2017). The DUP4 structural variant which encodes the Dantu blood group and is only found in East Africa, is known to protect against severe malaria through increased surface tension of the erythrocyte thereby inhibiting *P. falciparum* erythrocyte invasion (Kariuki et al., 2020). Higher haemoglobin levels have also been linked to DUP4, suggesting that DUP4 protects against severe malaria anaemia (Algady et al., 2018). Moreover, data from functional studies have demonstrated partial resistance to *P. falciparum* invasion in GPA and GPB deficient RBCs (Hadley et al., 1987), and that a GPC allele, highly prevalent in Melanesians inhibits *P. falciparum* invasion (Maier et al., 2003). However, recent functional studies in different populations are yet to confirm these findings.

Unlike the DUP4 variant that has been well characterised and its protective role with malaria severity established, the relationship between malaria and other glycophorin variants like the deletions (especially the most commomn GPB deletions 1 and 2) highly prevalent in West Africa is yet to be established. However, a study in Ghana observed that GPB deletion 1 and 2 alleles were associated with malaria-negative individuals (Amuzu et al unpublished data).

2.5.4 Methods of Genotyping Glycophorin Gene Variants

The glycophorin genes (GPA, GPB, and GPC) are evolutionarily related and share about 96 % sequence homology due to duplication and recombination events between GPA, and GPB (Leffler, et al., 2017; Polin et al., 2014). This high sequence homology makes it challenging to develop targeted and less expensive genotyping approaches to specifically study/characterise each variant and map their distribution in malaria-endemic areas. Hence, genome sequencing became the main approach to studying glycophorin variants. However, PCR genotyping

approaches still play a significant role especially for studies with larger sample sizes and in areas with limited resources (Hollox & Louzada, 2023). Thankfully, with technological advancements and high research interest in the glycophorin locus, researchers have been able to break the high sequence homology barrier and PCR-based genotyping methods have been developed for some of the variants. For instance, PCR-based restriction fragment length polymorphism (RFLP), paralogue ratio, and breakpoint-specific PCR protocols have been developed and are currently being used for genotyping of GPB variants (Algady et al., 2021; Amuzu et al., 2021; Lane et al., 2020). PCR-RFLP and breakpoint-specific PCR approaches have also been developed for the genotyping of DUP4 (duplication of GPA and GPB regions) (Algady et al., 2018; Leffler, et al., 2017). The development of these genotyping protocols is timely in this era where research is focus towards malaria elimination with genomic surveillance being the key approach. Additionally, these approaches provide a cheaper and faster way of mapping the distribution of these variants in malaria-endemic areas and determine their associations with malaria outcomes which is one of the objective of this study.

Nonetheless, there are still challenges in developing SNP and structural variation genotyping approaches as accuracy can be limited by incorrectly mapping short sequencing reads between paralogues, especially in areas where gene conversion alleles have occurred. Thus, for some of the structural variants, long-read sequencing is still the recommended genotyping approach for increased reliability and improved haplotype phasing (Hollox & Louzada, 2023).

CHAPTER 3

Recent Increase in Low Complexity Polygenomic Infections and Sialic Acid-Independent Invasion Pathways in *Plasmodium falciparum* from Western Gambia

3.1 Abstract

The malaria parasite *Plasmodium falciparum* utilizes multiple alternative receptor-ligand interactions for the invasion of human erythrocytes. While some *P. falciparum* clones make use of sialic acid (SA) residues on the surface of the human glycophorin receptors to invade the erythrocyte, others use alternative receptors independent of sialic acid residues. In this study, it was hypothesised that "erythrocyte invasion pathways of parasites in The Gambia will be dominated by sialic-acid independent mechanisms driven by the expression of specific ligands".

Blood samples were collected from 90 malaria-infected participants with uncomplicated malaria across four years (2015, 2016, 2021, and 2022) in The Gambia. Genetic diversity was determined by genotyping the merozoite surface protein 2 (MSP2) polymorphic gene of *P. falciparum.* Erythrocyte invasion phenotypes were determined using three enzymes: neuraminidase, trypsin, and chymotrypsin, known to cleave different receptors from the surface of the erythrocyte. Schizont-stage transcript levels were obtained for a panel of six *P. falciparum* invasion ligand genes (*eba175, eba181, Rh2b, Rh4, Rh5, and clag2*) using 48 isolates with quality RNA. The multiplicity of infection (MOI) was determined by dividing the total number of alleles identified for MSP2 gene by the total number of samples genotyped while the Nei unbiased expected heterozygosity (H*e*) was used to estimate genetic diversity. The Shapiro-Wilk normality test was performed to determine if the distribution of the data was Gaussian, and variables that passed the test were subjected to parametric analysis; otherwise, non-parametric techniques were employed.

Though the allelic heterozygosity of MSP2 repeat region decreased as expected with reduced transmission, there was an increase in infections with more than a single MSP2 allelotype from 2015 to 2022. The invasion phenotypes of these isolates were mostly SA-independent with a continuous increase from 2015 to 2022. Isolates from 2021 and 2022 were highly inhibited by chymotrypsin treatment compared to isolates from 2015 and 2016. Higher invasion inhibition for 2021 and 2022 isolates were further obtained following erythrocyte treatment with a combination of chymotrypsin and trypsin. The transcript levels of invasion ligand genes varied across years. However, levels of *clag2,* a rhoptry-associated protein, were higher in 2015 and 2016 isolates than in 2021 isolates, while the *Rh5* transcript level was higher in 2021 compared to other years.

Overall, these findings suggest increasing polyclonal infections with an increase in the use of sialic-acid independent invasion pathways by *P. falciparum* clinical isolates in The Gambia.

3.2 Introduction

Plasmodium falciparum malaria parasites invade and multiply in human red blood cells (RBCs) during the entire erythrocytic phase of its life cycle. Erythrocyte invasion by *P. falciparum* merozoite is a crucial, complex, and multistep process requiring multiple interactions between hot cells and the parasite (Gaur et al., 2004). During this process the merozote surface proteins (MSPs) like MSP2 interacts with members of two major protein families, including reticulocyte-binding homologues (Rhs) and erythrocyte-binding antigens (EBAs). The EBAs protein family members used for erythrocyte invasion includes EBA-175, EBA-140, EBA-165, EBL1, and EBA-181 (Gilberger et al., 2003; Mayer et al., 2009; Narum et al., 2002) while the Rhs protein family members are Rh1, Rh2a, Rh2b, Rh4, and Rh5 (DeSimone et al., 2009; Rayner et al., 2001; Tham et al., 2009; Triglia et al., 2001). MSP2 found on the surface of the merozoite is a very polymorphic gene used in genetic diversity studies (Fenton et al., 1991) which also plays an role in the preliminary attachment of the parasite to the surface

of the erythrocyte during invasion. This attachment and surface localisation is critical for the parasite's ability to recognise and penetrate the erythrocytes. Following attachement, EBAs and Rhs protein families interacts with erythrocyte receptors to enable a successful invasion (Molina-Franky et al., 2022). These proteins determines the invasion phenotypes of *P. falciparum*, which have mostly been characterized by the treatment of erythrocytes with enzymes known to cleave part of the receptor repertoire on which they bind. Commonly used enzymes include neuraminidase which cleaves sialic acids from the glycophorin (GP) receptors (GPA, B, and C), trypsin cleaves glycophorins A, C, and complement receptor 1 (CR1), and chymotrypsin cleaves GPB, band 3 and other invasion related proteins (Cowman & Crabb, 2006; Triglia et al., 2011). Thus, based on the sialic acid (SA) residues of the glycophorin receptors, *P. falciparum* invasion pathways are either described as SA-dependent or SA-independent. These ligand-receptor interactions are variably used by *P. falciparum* isolates in different malaria populations and some parasite lines activate alternative pathways or switch invasion phenotypes during *in vitro* life cycle (Hadley et al., 1987; Okoyeh et al., 1999; Perkins & Holt, 1988).

Previous studies in The Gambia have reported the predominant use of the SA-dependent pathways during a period of relatively high malaria transmission (Baum et al., 2003; Gomez-Escobar et al., 2010). Over the decade, malaria transmission has drastically reduced due to intensified and sustained malaria control interventions. Currently, The Gambia is moving towards malaria elimination (pre-elimination phase) with an overall prevalence of 0.2 % (WHO, 2022a). Hence, it is not known how this transition from high to low malaria transmission has affected erythrocyte invasion mechanisms, which involves ligands targeted for vaccine development.

In this study, the invasion phenotypes and transcript levels of six common invasion ligand genes (*eba175, eba181, Rh2b, Rh4, Rh5, and clag2*) of *P. falciparum* field isolates from four

different years in The Gambia (2015, 2016, 2021, and 2022) was investigated. The highly polymorphic *P. falciparum* surface protein, merozoite surface protein 2 (MSP2) was also used to assess the complexity of infection from clinical isolates across these years.

3.3 Methods

3.3.1 Ethical Approval

This study made used of samples collected from a previous project in 2015 and 2016 with permission to reuse them and samples collected in 2021 and 2022 (the PhD duration). For the 2021 and 2022 samples, Ethical clearances for both projects were obtained from The Gambia Government/Medical Research Council Unit in The Gambia (MRCG) Joint Ethics Committee as shown on appendix I. The research protocol was explained to the participants in both English and a local language for better understanding and a signed informed consent (appendix II) was sought before enrolment. A structural questionnaire was used to collect demographic information from consented participants (appendix III).

3.3.2 Sample Size Determination

All participants for this project were recruited from the same sites/health facilities and each project objective had a different sample size. The sample sizes were determined per objective depending on the hypothesis and the complexity of the experimental approach used. For this objective, the sample size calculation was done using the Cochran's sample size formula (Cochran, 1977) shown bellow, taking into consideration the prevalence of malaria in The Gambia (0.2 %) and a margin of error of 1 %. The Z-value for a level of confidence of 95 % (1.96) was used to determine an initial sample size and then it was corrected for a finite number of 2.8 million people which is the population of The Gambia. This calculation yielded a final sample size of 77 people.

$n = \frac{Z^2.\ p.\ (1-p)}{E^2}$ Where: Z is the Z-value for a 95% confidence level (1.96), p is the prevalence of malaria in The Gambia, and E is the margin of error.

3.3.3 Sampling Sites and Study Population

This was a hospital-based study and convinience sampling approach was used to recruit participants. Participants were recruited from four health facilities in three regions of the country. These included the Brikama and Fajikunda health facilities in the West Coast Region (WCR), Edward Francis Small Teaching Hospital (EFSTH) in the Greater Banjul Region, and the Basse Health centre in the Upper River Region (Figure 3.1). Recruitment was done between September and December in the years 2015, 2016, 2021, and 2022. The samples from 2015 and 2016 were obtained from a prior project that included preserved *P. falciparum* isolates. These preserved samples were utilised for this study to facilitate culture-based analyses. In contrast, the samples collected during 2021 and 2022 were gathered specifically for this PhD research. There were no preserved *P. falciparum* isolates available for the intervening years of 2017 through 2020 to be cultured and used for invasion and or gene expression analysis. Consequently, samples from these years were excluded from this part of the study since a convinient sampling approach was used.

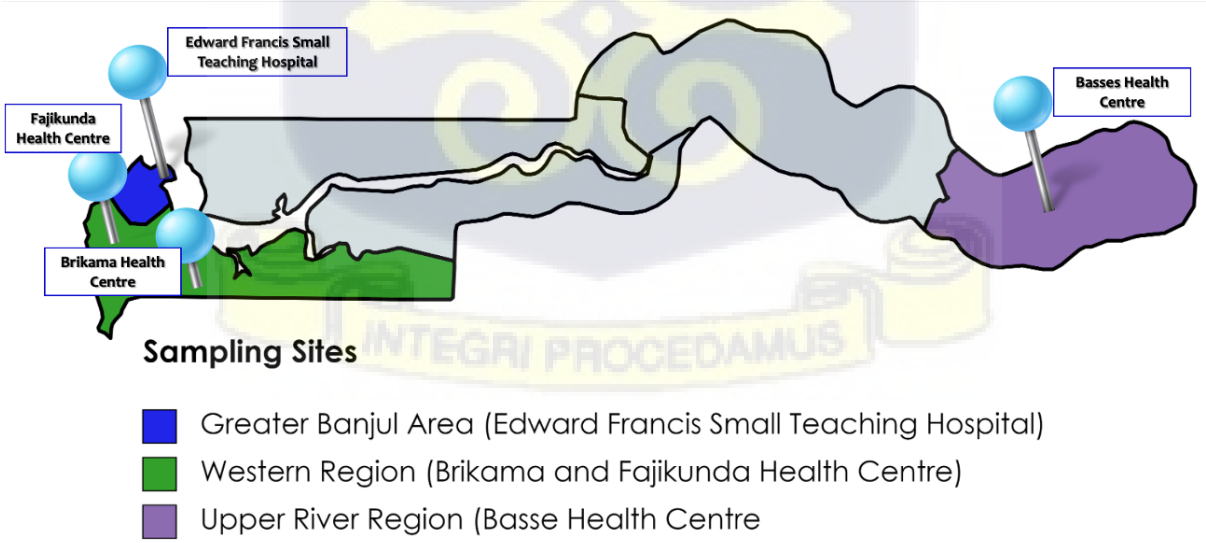


Figure 3.1: Map of The Gambia showing the sampling sites

3.3.4 Inclusion/Exclusion Criteria

The inclusion criteria for this project was people with fever (axillary temperature ≥ 37.5 °C), a history of fever and/or any other sign and symptom of malaria in the last 48 hours who consulted at any of the health facilities between September to December (malaria season) of each sampling year and consented to participate. Those who tested positive for *P. falciparum* malaria by immunochromatic rapid diagnostic testing (RDT) and reported not to have taken antimalarial drugs during the preceding three days were also invited to participate in the study. There was no age group nor gender restrictions in the recruitment procedures.

3.3.5 Sample Collection and Processing

After consenting, 2 ml of whole blood was collected from each participant into labeled EDTA tubes and transported in a cold box to the malaria research laboratory at MRCG within 4 hours for further processing. At the laboratory, approximately 50 µl of the blood was spotted in a circular pattern on a filter paper (Whatman 3 mm, GE Healthcare, USA) labeled with the participant's identification number and allowed to air dry overnight. Once dried, the dry blood spots (DBS) were packaged individually in ziploc bags with desiccant, sealed, and stored at 4 °C until used. Two *P. falciparum* laboratory strains 3D7 and Dd2 were used as controls for all experiments.

For each participant, a thin blood film was prepared, fixed with methanol for a minute, stained with 10 % Giemsa solution for 10 minutes, and examined under oil immersion (x100 magnification) of a conventional light microscope. Parasitaemia was estimated using the standard method of counting the number of asexual parasites per 1000 RBCs. All positive samples by microscopy were then processed for DNA extraction, parasite culture and invasion experiments. A total of 90 malaria positive samples were used for this objective.

3.3.6 DNA Extraction

Genomic DNA (gDNA) was extracted from whole blood alongside a laboratory strain (3D7) as positive control using the QIAamp® DNA Blood Mini kit (Qiagen, Cat 51104) following the manufacturer's protocol. The concentration of the extracted gDNA was determined using the Qubit dsDNA high-sensitivity kit (Thermo Fisher Scientific, Cat Q32854). To avoid repeated freeze-thaw cycles and degradation, the extracted DNA was divided into 2 aliquots (100 µl each), and the aliquot for immediate use was stored at -20°C while the second aliquot was labeled as stock and stored at -80°C.

3.3.7 Genotyping of Merozoite Surface Protein 2 (MSP2)

Using family-specific primers, nested PCR was performed targeting the core polymorphic regions of MSP2 for two *P. falciparum* allelic families (FC27 and IC3D7) (Snounou et al., 1999). The primary PCR reaction was done in a final volume of 15 µl, and the nested PCR in a final volume of 20 µl. Both primary and nested PCR constituents and cycling conditions are shown in appendix 5, table S1. The 3D7 and Dd2 *P. falciparum* laboratory strains were used as PCR-positive controls while nuclease-free water was used as negative control. QIAxcel (Qiagen) and QIAxcel ScreenGel Software were used to separate nested PCR products and determine fragment sizes respectively (appendix 5, Figure S1 A and B).

3.3.8 Parasite Culture, Schizont Enrichment, and RNA Extraction

A total of 90 clinical isolates with sufficient volumes (500 µl) and parasitaemia not less than 0.3 % were cultured at 2-5 % haematocrit for a maximum of 48 hours in RPMI 1640 medium (Sigma-Aldrich, Cat 27016-021) enriched with 0.5 % Albumax II (Gibco-BRL, Cat 11021-037), 35 mM HEPES solution (Sigma-Aldrich, Cat H0887-100ML), 1 mg/litre of hypoxantine (Sigma-Aldrich, Cat H-9636), 45 % D-glucose solution (Gibco-BRL, Cat G8769-100ML), 3 mg/ml L-cysteine (Sigma-Aldrich, Cat C7352-25G), L-glutamine solution (Sigma-Aldrich, Cat G7513-100ML), and 5 µg/ml of gentamicin solution (Gibco-BRL, Cat G1272-100ML).

The culture flasks were gassed with a mixture of 1 % O_2, 3 % CO_2, and 90 % N_2 and incubated at 37 ºC. Culture media were changed every 24 hours following microscopic examination of the parasite's stage and to estimate to appropriate time to arrest the culture for schizont harvesting. All samples were maintained in culture under the same conditions with constant monitoring until parasites were at the schizont stage of development (appendix V, Figure S2A). The 3D7 laboratory-adapted strain was equally maintained in culture under the same conditions as the clinical isolates, harvested at the schizont stage, and used as positive control for all downstream experiments. The schizonts-rich erythrocytes were resuspended in three parts of TRIzol® reagent (Ambion, Cat 15596018) and stored at -80ºC until RNA extraction as previously described (Gomez-Escobar et al., 2010). The Phenol-chloroform RNA extraction protocol (Appendix IV) was used for RNA extraction and the Qubit High Sensitivity RNA kit (Thermo Fisher Scientific, Cat Q32852) was used for quantification. RNA was stored at -80ºC for short-term use and liquid nitrogen for long-term use.

3.3.9 Reverse Transcriptase Quantitative PCR (RT-qPCR)

RNA from each isolate was thawed, reverse transcribed using the ProtoScript® First Strand cDNA Synthesis Kit (New England Biolabs, Cat E6300L), and the transcript levels of six invasion ligand genes: *eba175, eba181, clag2, Rh2b, Rh4*, and *Rh5* were determined using gene-specific primer/probe sets from previous studies (Blair, 2002; Gomez-Escobar et al., 2010; Nery et al., 2006) (appendix V, Table S2). Real-time quantitative PCRs were performed using the TaqMan universal PCR Master Mix (Thermo Fisher Scientific, Cat 4326708) in 15 µl volumes with 330 nM concentrations of each primer and 160 nM concentrations of each probe following the cycling conditions indicated in appendix V, table S2. Each run included controls and 3D7 genomic DNA standards, with standard curves generated for each run (appendix V: Figure S3). The threshold fluorescence value for each run was determined automatically by the BioRad CFX96 software.

3.3.10 Treatments and Labelling of Target Erythrocytes

Targeted uninfected RBCs were donated by laboratory volunteers after screening to ensure they were malaria-free. Upon confirmation of their malaria status, a blood grouping test was done and only RBCs from malaria-negative blood group O^+ individuals were used for downstream experiments. After leukocyte depletion, the targeted erythrocytes were treated with three different enzymes: 2 UM/ml neuraminidase (Sigma Aldrich, Cat N6514), trypsin (Sigma Aldrich, Cat T9935) and chymotrypsin (Sigma Aldrich, Cat C4129) in six different combinations and a negative control (RPMI only) as shown on Table 3.1 below. Enzyme treatment was done for 1 hour on a rotator at 37 °C. To stop the treatment process, erythrocytes were washed three times with 1X PBS (Sigma-Aldrich, Cat SLCL7026), resuspended with iRPMI to 2 % haematocrit (HCT) and stored at 4 °C for a maximum of 24 hours before labelling.

The treated erythrocytes were labelled with the Cell Trace Far Red (CTFR) fluorescent dye (Invitrogen, Cat C34564), highly sensitive for labelling erythrocytes and commonly used for flow cytometry-based proliferation assays (Filby et al., 2015; Theron et al., 2018). A previously optimised concentration of 10 µM was used and staining was done at 2 % HCT cells suspension in iRPMI for 2 hours at 37 °C on a rotator mixer. To stop the labelling reaction, cells were incubated with iRPMI for 30 minutes after which they were washed three times with phosphate buffer saline (PBS). The enzymatically treated and labelled targeted erythrocytes were then resuspended at 2 % HCT with complete RPMI for invasion assay set up.

Table 3.1: Enzyme combinations and concentrations

Enzyme combinations	Volume of enzyme (Stock Concentration	Volume of iRPMI	Volume of Target erythrocyte	Final Concentration
Neuraminidase (Nm)	13.3 µl	120 µl	66.67 µl	66.7 mU/ml
Low Trypsin (LT)	4 µl	129.3 µl	66.67 µl	66.7 µg/ml
High Trypsin (HT)	60 µl	73.3 µl	66.67 µl	1.0 mg/ml
Chymotrypsin/ Low trypsin (CHY_LT)	60 µl Chymotrypsin, 4 µl trypsin	69.3 µl	66.67 µl	1.0 mg/ml Chymotrypsin, 66.7 µg/ml trypsin
Chymotrypsin (CHY)	60 µl	73.3 µl	66.67 µl	1.0 mg/ml
All (Positive control)	13.3 µl Nm, 60 µl chymotrypsin, 60 µl trypsin	0 µl	66.67 µl	66.7 mU/ml Nm, 1.0 mg/ml trypsin, 1.0 mg/ml chymotrypsin
RPMI (Negative control)	0 µl	133.3 µl	66.67 µl	NA

3.3.11 Invasion Assay Set up

For invasion set up, aliquot of ring stage (appendix V, Figure S2B) infected red blood cells (iRBCs) from the field (unculture samples) was diluted with complete RPMI to 2 % haematocrit. Isolates with initial parasitaemia greater than 0.3 % were diluted to 0.3 % with uninfected red blood cells from volunteer blood group O^+ donors while isolates with parasitaemia less than or equall to 0.3 % were not deleted. 25 µl of the sample was transferred to each well of a flat-bottom 96-well plate according to the plate set up (Figure 3.2). 75 µl of the enymattically treated CTFR labelled erythrocytes were then added to the appropriate wells resulting in a total culture volume of 100 µl per well. A smear well (with no enzyme treated labelled erythrocytes) was added per sample and used for thin film preparation prior to assay harvesting after 48 hours to monitor and ensure complete reinvasion. All invasion assays were carried out in triplicates and each well suspension was thoroughly mixed by repeated pipetting.

The plate was then placed in a gas chamber and a mixture of gas (1 % O_2, 3 % CO_2, and 96 % N) was run through the chamber followed by sealing and incubation at 37 °C for 48 hours. After 48 hours, the plate was removed from the incubator and using the smear wells, a thin film was made, stained with 10 % geimsa for 10 minutes and checked with a light microscope to confirm complete reinvasion through observation of early rings before assay harvesting. Parasites labelling was achieved by adding 200 µl of 1:5000 SYBR Green 1 nucleic acid gel stain (Invitrogen, S7563) in PBS to each well and the plate was incubated at 37 °C for 20 minutes. Following incubation, plates were centrifuged at 1500 rpm for 5 minutes and each well was washed with 200 µl PBS. The PBS wash was repeated twice before each pellet was resuspended with 200 µl PBS. 50 µl aliquots were transferred from each well into a new flat-bottomed 96-well plate containing 150 µl and examined on a flow cytometer (BD Accuri C6 Plus Sampler) within 24 hours and the population of interest (heterologous reinvasion) was gated (Figure 3.2). A total of 100,000 events were acquired for each well, and data were analysed using the BD Accuri C6 Plus Sampler software.

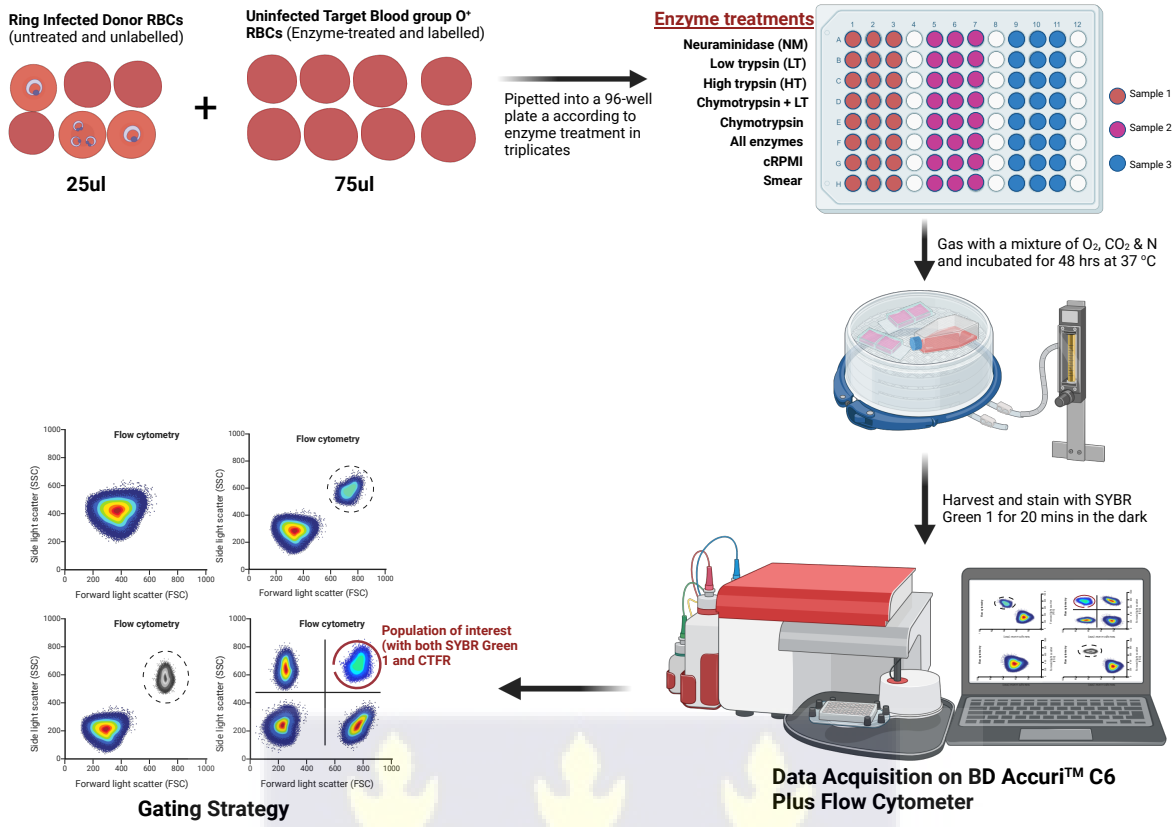


Figure 3.2: Flowchart of invasion assay setup (Drwan with biorender).

3.3.12 Data Analysis

The multiplicity of infection (MOI) was determined by dividing the total number of alleles identified for MSP2 gene by the total number of samples genotyped (Mayengue et al., 2005), while the Nei unbiased expected heterozygosity (H*e*) was used to estimate genetic diversity. H*e* provides an indication of the probability that alleles drawn from two individuals will be different. It ranges from zero (no genetic diversity) to one (high genetic diversity). H*e* was estimated using the formula, $He = [n/(n-1)]\ [(1-\sum Pi^2)]$, where n is the number of samples and Pi is the frequency of alleles in each locus (Nei, 1978).

R version 4.2.0 was used to analyse data related to invasion and gene expression. The Shapiro-Wilk normality test was performed to determine if the distribution of the data was Gaussian, and variables that passed the test were subjected to parametric analysis; otherwise, non-parametric techniques were employed. Using the Kruskal-Wallis test, relative gene expressions

and enzyme inhibition rates were compared across sample years. Post hoc pairwise comparisons were then carried out using the Mann-Whitney U test with Bonferroni adjustment for multiple pairwise comparisons. The Spearman's (rho) method was used to investigate correlations between the enzyme inhibition phenotypes. For all analyses, P values ≤ 0.05 were considered statistically significant.

3.4 Results

3.4.1 Demographic and Parasitological Information of Study Participants

A total of 90 malaria positive samples were used with 57 (63.3 %) from Brikama (Western region), 17 (18.9 %) from Fajikunda (Western region), 14 (15.6 %) from Basse (Upper River region), and 2 (2.2 %) from EFSTH (Greater Banjul region). Based on the recruitment years, the lowest proportion came from 2015 (16.67 %) while 2016, 2021, and 2022 had similar proportions (28.89 %, 26.67 %, and 27.78 % respectively). While the mean age was similar between the 2015 and 2016 participants (9 and 8.21 respectively), it was relatively higher for 2021 (20.48) and 2022 (16.67) participants. This is because only children of 0-15 years were recruited in 2015 and 2016 whereas 2021 and 2022 participants were from all age groups (1 - 70 years). The proportion of female participants 46 (51.1 %) was higher than that of male participants 36 (40 %) and 8 (8.9 %) participants did not indicate their gender. Samples collected in 2015 had higher parasitaemia ranging from 1 to 7 % with a median of 1.85 % compared to 2016, 2021, and 2022, indicating a reduction in parasitaemia across the years from 2015 to 2022, however, these differences were not statistically significant (P = 0.213) across the years (Table 3.2 and Appendix V: Figure S4).

Table 3.2: Demographic characteristics of study participants across four years (n =90)

Variables	2015	2016	2021	2022
Number (%)	15 (16.67)	26 (28.89)	24 (26.67)	25 (27.78)
Mean age ± SD	9 ± 3.74	8.21 ± 3.44	20.48 ± 10.28	16.67± 9.03
Male (%)	8 (61.5)	10 (41.7)	13 (54.2)	5 (23.8)
Female (%)	5 (38.5)	14 (58.3)	11 (45.8)	16 (76.2)
Median Parasitaemia (range)	1.85 (1-7 %)	1.24 (0.21-5.2 %)	1(0.4-3.6 %)	0.9 (0.1-6.4)

SD = Standard deviation

3.4.2 Genetic Diversity of *P. falciparum* Infection in The Gambia

MSP2 genotyping identified a total of 32 individual alleles, 22 from 3D7 allelic family (fragment range 230-700 bp) (Figure 3.3A) and 10 from FC27 allelic family (fragment range 250-780 bp) (Figure 3.3B). The frequencies of all alleles detected in the 3D7 family were less than 20 % and a 350 bp allele was the most prevalent in both 3D7 and FC27 families (13 % and 25 % respectively).

The frequency of likely monogenomic infections with a single 3D7 allele was higher in 2016 (61.5 %) isolates compared to 2015 (40 %), 2021 (40.9 %), and 2022 (21.7 %) isolates. However, these differences were not statistically significant (P = 0.39). Using a 95 % confidence level, the confidence interval for these data was (22.13, 63.66). Furthermore, there was an increase in polygenomic infections of the same 3D7 allelotype in 2022 (30.4 %), compared to 2021 (13.6 %), 2015 (6.6 %), and 2016 (7.7 %), but these differences were also not statistically significant (P = 0.4). Somewhat similarly, isolates with the FC27 allelotype had a higher frequency (33.3 %) of likely monogenomic infections in 2015 with a confidence interval between 9.3 and 34.5. In the analysis of polygenomic infections with both 3D7 and FC27 allelic types, a notable increase in prevalence was observed in the 2022 isolates, reaching 68.9 %. This represents a significant rise compared to previous years: 36.6 % in 2021, 20 % in

2015, and 15.4 % in 2016. Using a 95 % confidence level, the confidence interval for these data is estimated to be between 13.3 and 56.6% with a P value of 0.4 (Figure 3.3C).

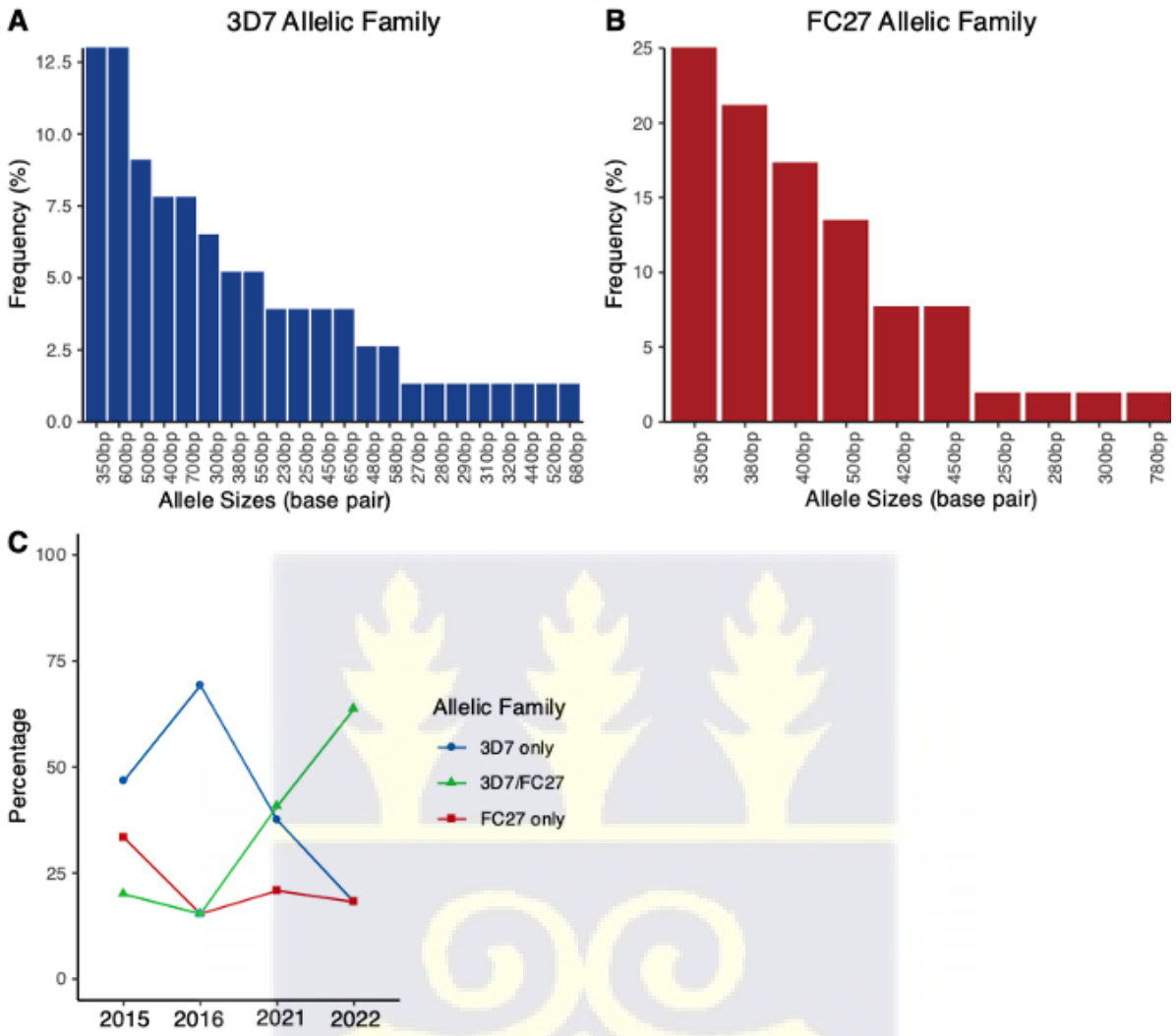


Figure 3.3: Prevalence and allelic frequencies of MSP2. A) frequencies of the 3D7 allelic family, B) frequencies of the FC27 allelic family, and C) Prevalence of MSP2 repeat polymorphism allelic families of *P. falciparum* in clinical isolates. bp = base pairs

3.4.3 Multiplicity of Infection (MOI) and Heterozygosity (*He*)

The presence of multiple clones in a single infection as defined by the MOI index was higher in 2022 (2.0) than in 2021 (1.55), 2016 (1.27), and 2015 (1.33) isolates. There was an overall decrease in the heterozygosity (*He*) of the MSP2 gene from 2015 (0.29) through 2016 (0.2),

2021 (0.06) to 2022 (0.34) (Table 3.3). However, the differences in both MOI and *He* across the years were not statistically significant (P = 0.4 and 0.39 respectively).

Table 3.3: MOI and *He* of MSP2 gene of *P. falciparum* across four years.

Year	MOI	*He*
2015	1.33	0.29
2016	1.27	0.2
2021	1.55	0.06
2022	2.0	0.34

3.4.4 Gene Expression of *P. falciparum* Invasion Ligands

Specific merozoite ligand genes are predominately expressed by *P. falciparum* isolates during the invasion process and could determine the pathways used. To determine variation in transcript levels of six invasion ligands (CLAG-2, EBA-175, EBA-181, RH-2B, RH4, and RH5), each isolate was cultured *ex vivo* to predominantly schizont stages and bulk RNA was extracted. A total of 48 out of 90 isolates cultured were successfully arrested at mature schizont stage within the first 48 hours of culture for transcript analysis [2015 (n =12), 2016 (n =15), 2021 (n =21), and 2022 (n = 0)]. For the other 42 isolates, data was not available either due to poor growth in culture or too low RNA yield after extraction to allow for reliable transcript quantification. The combined transcript profiles across all ligands grouped the isolates into four main clusters (A, B, C, and D). Transcript level profiles of the isolates clustered RH4 and RH-2B as well as RH5 and RH-2B together, which are associated with sialic acid-independent invasion pathway, while EBA-181 and EBA-175 genes associated with sialic acid-dependent invasion pathway also clustered together (Figure. 3.4A).

Further correlation analysis between the genes showed significantly positive correlations between EBA-175 and EBA-181 (Spearman's $r = 0.5$, $P = 2e{-}04$), RH-2B and RH4 ($r = 0.5$, $P = 0.001$), and RH5 and RH4 (Spearman's $r = 0.7$, $P = 1e{-}06$) (Figure. 3.4B).

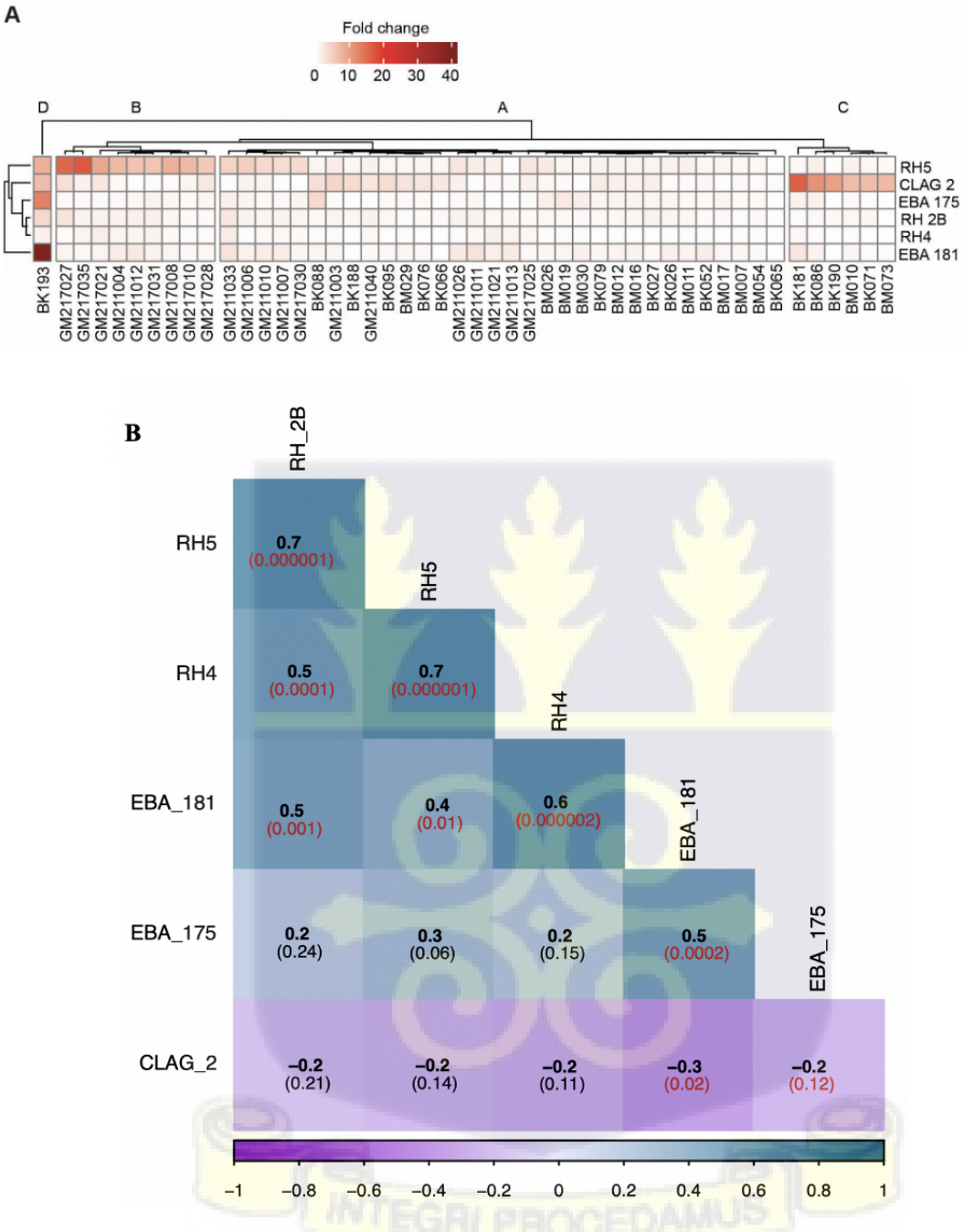


Figure 3.4: Relative expression and correlation between expression levels of invasion ligand genes. A) Hierarchical clustering heatmap of relative expression levels of six *Plasmodium falciparum* ligand genes (rows) between clinical isolates (columns). White to light red colour signifies lower expression while brown colour indicates higher

expression. B) Correlation between gene expression levels; numbers out of the brackets in square boxes represent Spearman's correlation coefficients while those in brackets represent the *P* values. Significant *P* values are indicated in red.

3.4.5 Expression of Invasion Ligand Genes of *P. falciparum* Clinical Isolates Across Sampling Years

CLAG2 (mean 2.61%) and RH5 (mean 2.93 %) were the most abundant genes expressed with no significant difference between their expression ($P = 0.41$), followed by EBA-181 (mean 1.05 %) and EBA-175 (mean 1.03 %) while RH4 was the least expressed with a mean of 0.43 (Figure 3.5A). When compared across the different sampling years, the expression of CLAG2 was higher in parasites collected in 2015 than RH5 ($P = 0.007$) (Figure 3.5B) unlike in 2016 where there was no significant difference between CLAG2 and RH5 ($P = 0.55$) (Figure 3.5C). Moreover, RH5 was highly expressed in 2021 isolates than CLAG2 and these differences were statistically significant ($P < 0.001$) (Figure 3.5D).

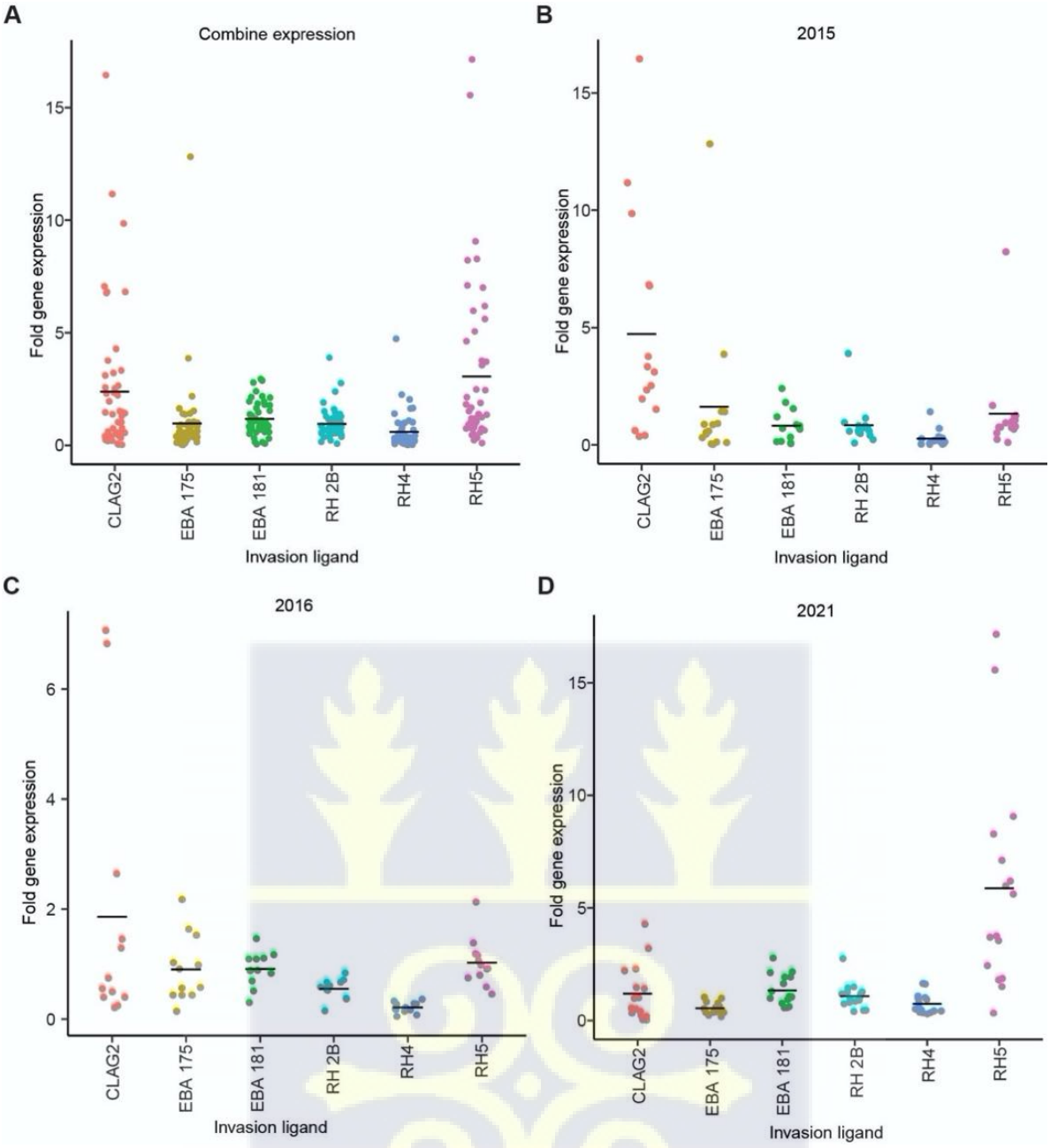


Figure 3.5: Fold gene expression of genes encoding six invasion ligands in *ex vivo* schizont-stage cultures of *P. falciparum* clinical isolates in The Gambia. (A) Combined expression of isolates across all years (n= 48), (B) 2015 isolates (n=12), (C) 2016 isolates (n=15), and (D) 2021 isolates (n= 21). Black horizontal lines indicate the mean for each transcript across all isolates sampled in each year and each dot denotes the transcript for each gene in a single clinical isolate.

3.4.6 Invasion Phenotypes of *P. falciparum* Across Four Years in The Gambia

Following cleavage of receptors using neuraminidase, trypsin, and chymotrypsin enzymes in different combinations, low levels (mean 29.83 %) of invasion inhibition into neuraminidase (NM) treated erythrocytes were observed in isolates from all four years compared to the level of inhibition by low trypsin (mean 46.23 %) but this was not statistically significant. Moreover, invasion inhibition by high trypsin (mean 83.49 %), chymotrypsin (mean 87.49 %), and chymotrypsin/low trypsin (mean 91.34 %) were all significantly higher than that of neuraminidase (all p < 2 e-16) (Figure 3.6). Though not statistically significant, it was noted that combining chymotrypsin and trypsin (CHY_LT) resulted in higher invasion inhibition (mean 91.34 %) than when the enzymes were used separately. This was expected because chymotrypsin and trypsin cleave different erythrocyte receptors. Hence combining them will result in the depletion of more receptors resulting in additive effect in invasion inhibition.

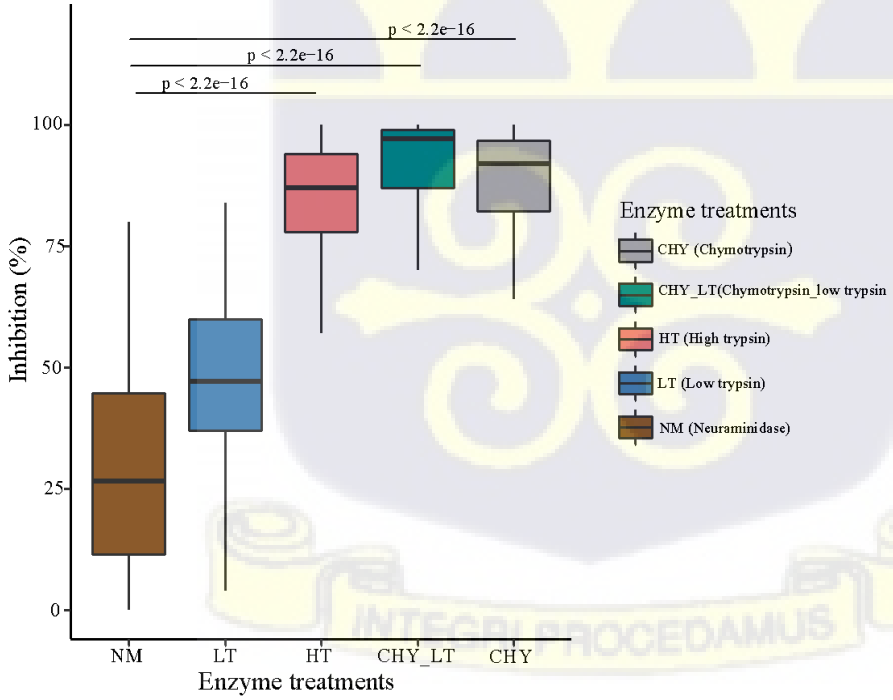


Figure 3.6: Invasion phenotypes of *P. falciparum* isolates across four years in The Gambia. Each box and whisker shows the distribution of inhibition, with the vertical line of each box representing the median. $P < 0.05$ were considered statistically significant.

3.4.7 Comparison of Invasion Phenotypes Across Sampling Years

Comparison of invasion phenotypes across the four years showed a trend towards increased neuraminidase inhibition between 2015 to 2022 (mean: 2015 = 22.21 %, 2016 = 26.69 %, 2021 = 31.6 %, and 2022 = 31.96 %) but the differences were not statistically significant ($P = 0.08$) (Figure 3.7A). Similar results were obtained for low trypsin (Figure 3.7B), high trypsin (Figure 3.7C), chymotrypsin/low trypsin treatments (Figure 3.7D) with increased enzyme inhibition across the years but with no significant differences ($P = 0.22$, 0.53, and 0.11, respectively). However, chymotrypsin treated erythrocytes resulted in significantly increased inhibition between 2015 and 2021 isolates ($P = 0.039$), 2015 and 2022 isolates ($P = 1.6e\text{-}05$), 2016 and 2022 isolates ($P = 0.001$), and between 2021 and 2022 isolates ($P = 0.024$) (Figure 3.7E).

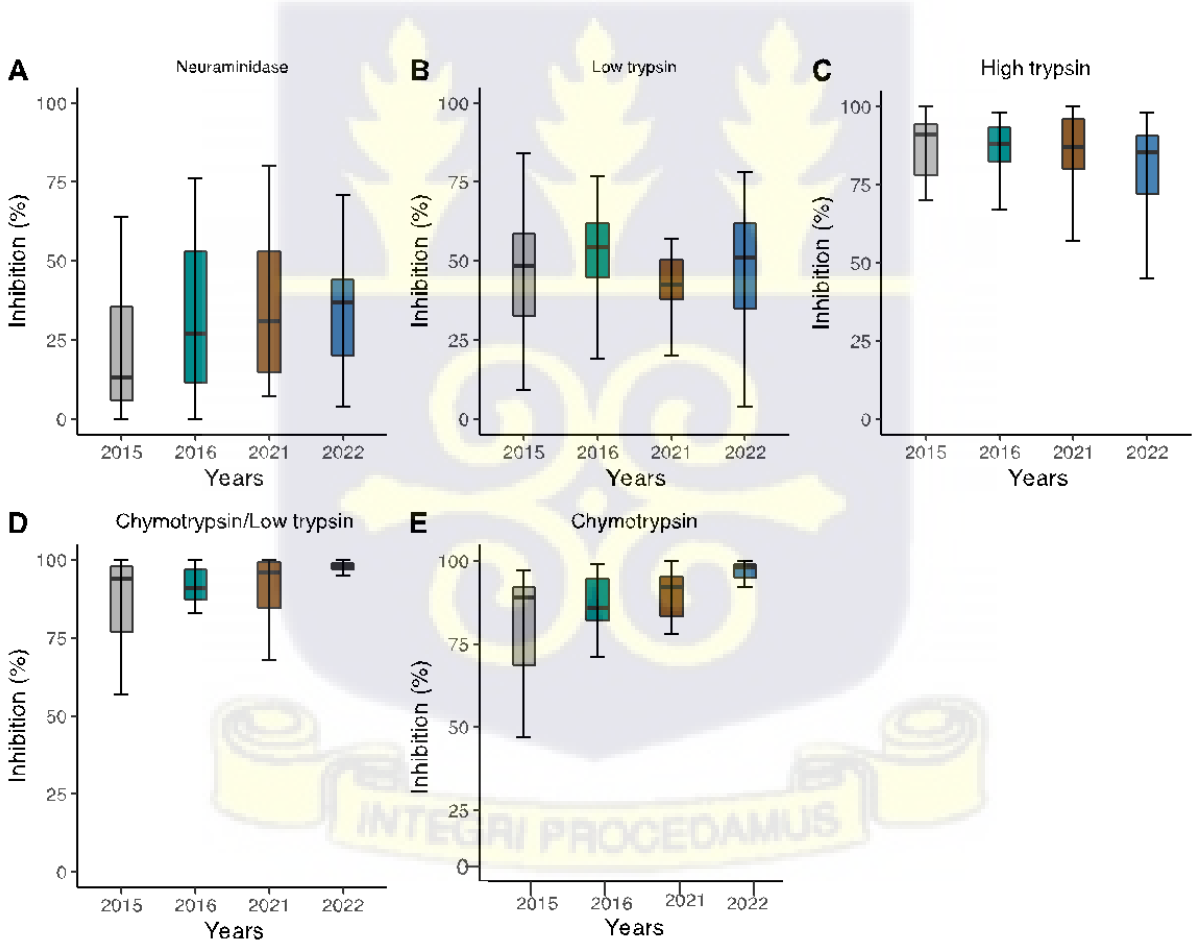
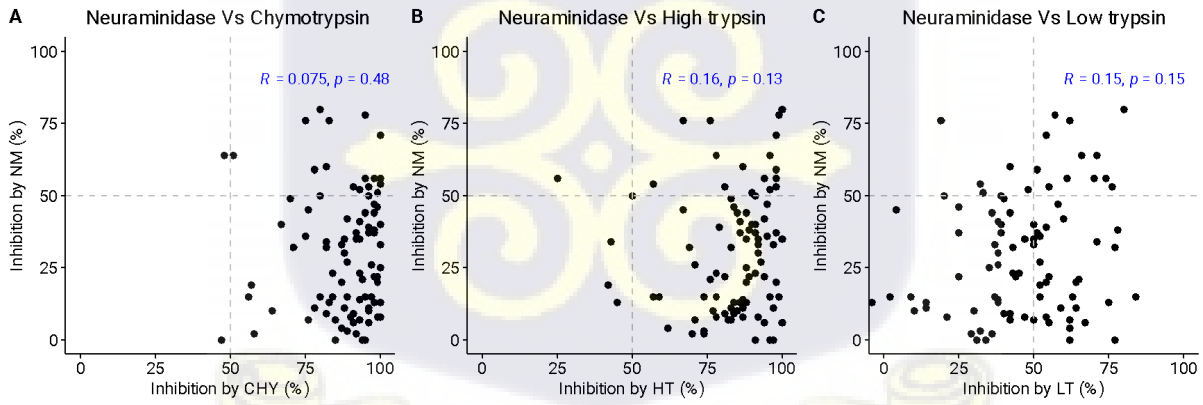


Figure 3.7: Comparison of erythrocyte invasion inhibition phenotypes of *P. falciparum* isolates across four years in The Gambia. A) neuraminidase, (B) low trypsin, (C) high trypsin (D) chymotrypsin/low trypsin, and (E) chymotrypsin. Each box shows the invasion

efficiencies in enzyme-treated erythrocytes relative to untreated erythrocytes and the horizontal black lines in each box represent the median. *P* values < 0.05 were considered statistically significant

3.4.8 Correlations between Invasion Inhibition by the Different Enzyme Treatments

Altogether, the inhibition of invasion of *P. falciparum* isolates by neuraminidase (NM) treatment of erythrocytes showed no correlation with inhibition by chymotrypsin (Spearman's R = 0.075, P = 0.48) (Figure 3.8A) and with inhibition by neuraminidase versus high trypsin treatment (Spearman's R = 0.16, P = 0.13) (Figure 3.8B), as well as inhibition by neuraminidase versus low trypsin treatment (Spearman's R = 0.15, P = 0.15) (Figure 3.8C). Invasion inhibition was however significantly positively correlated with neuraminidase versus chymotrypsin/low trypsin (CHY_LT) treatment (Spearman's R = 0.22, P = 0.034) (Figure 3.8D), chymotrypsin + high trypsin treatment (Spearman's R = 0.27, P = 0.009) (Figure 3.8E), and chymotrypsin + chymotrypsin/low trypsin (CHY_LT) treatment (Spearman's R = 0.5, P = 4.6e-07) (Figure 3.8F).



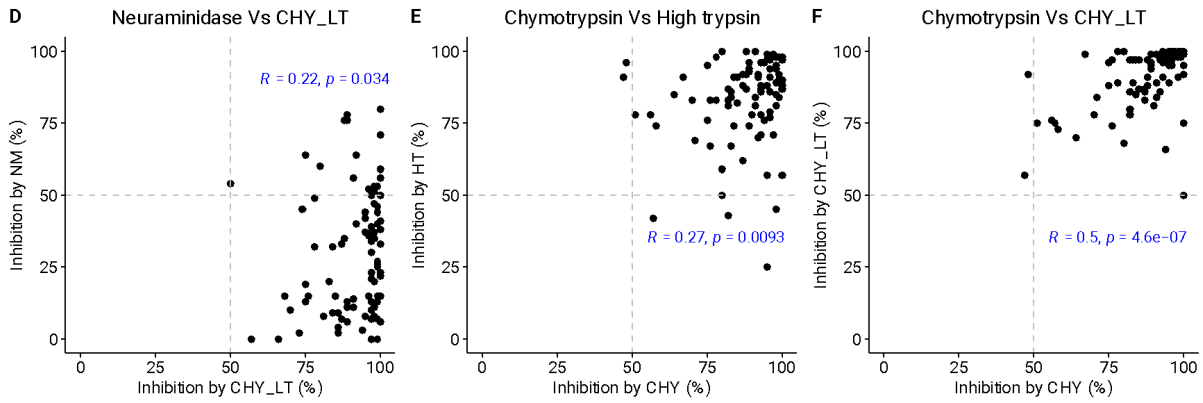


Figure 3.8: Pairwise correlations between invasion inhibition (%) by enzyme treatment. The Spearman correlation coefficient (*R*) and *P* values are shown in blue for each pair of comparisons.

3.4.9 Correlation Between Enzyme Treatments, Age Groups, and Parasitaemia

Since 2015 and 2016 isolates were collected from children < 15 years while 2021 and 2022 isolates were collected from all age groups (1 – 70 years), it was important to find out if different age groups and parasitaemia might have contributed to the observed differences apart from the sampling years. To this effect, the study participants were grouped into different age groups as follows: 0-5 years as young children, 6-17 years as older children, and 18 years and above as adults. Following this grouping, a multiple-variable correlation analysis incorporating the different age groups, parasitaemia, and the different enzyme treatments was carried out. Significantly negative correlations were observed between parasitaemia (PCT) and low trypsin (LT) treatment (R = -0.751, *P* = 0.045) among children while significantly positive correlations were observed between NM and LT treatments (R= 0.325, *P* =0.033), CHY and CHY_LT treatments (R = 0.627, P = 0.008) among older children. The correlation coefficient for each group and their corresponding statistically significant values are indicated in the upper triangle of the scatter matrix plot below (Figure 3.9).

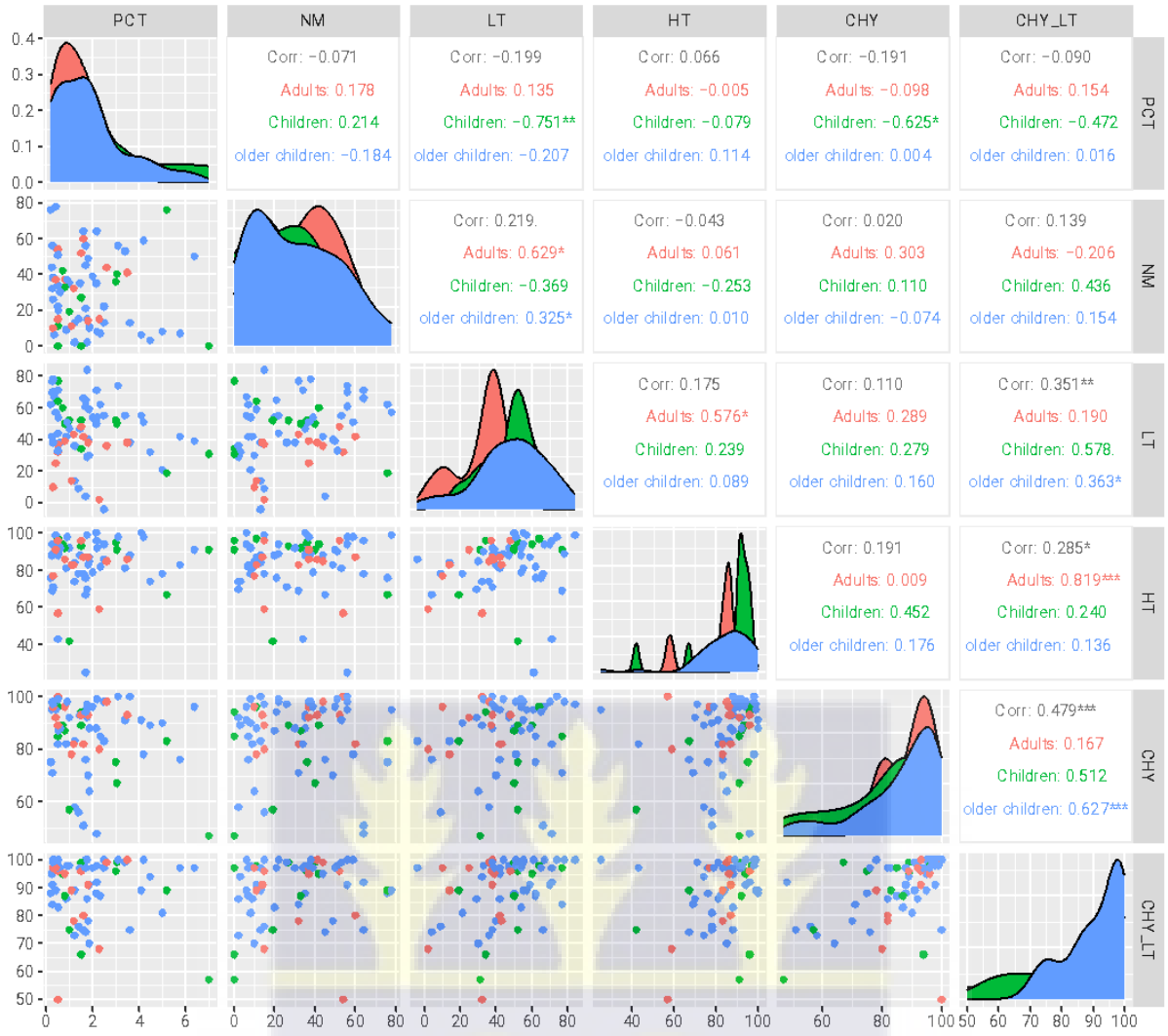


Figure 3.9: Scatter matrix plot of enzyme treatment, age groups, and parasitaemia between different age groups. Within each panel, each bullet point represents the percentage (%) invasion inhibition of each enzyme treatment. The colours represent the different age groups, while the asterisk (*) represents groups with significant *P*-values, NM = neuraminidase, LT = low trypsin, HT = high trypsin, CHY_LT = chymotrypsin/low trypsin, PCT = parasitaemia, and CHY = chymotrypsin.

3.4.10 Invasion Profile of *P. falciparum* Clinical Isolates According to Gene Expression Clusters and MSP2 Dominant Alleles

Further analysis of the invasion efficiency of isolates according to the ligand genes expression clusters in figure 3.4A of section 3.4.4 showed no statistically significant differences in invasion efficiency across all clusters within enzyme treatments (P = 0.46, 0.89, 0.23, 0.95, and 0.32 for NM, LT, HT, CHY_LT, and CHY respectively) (Figure 3.10A).

Following the observed variations in invasion inhibition phenotypes and gene expression across the years, it was important to test if the phenotypes also vary by MSP2 allelic families. For the most dominant allelic type (350Nbp) for both families (3D7 and FC27), invasion inhibition was significantly higher in the 3D7 allelic family only in the case of erythrocyte treatment with high trypsin (P = 0.027) (Figure 3.10B). For neuraminidase, low trypsin, and chymotrypsin-treated cells the 3D7 type isolates were less inhibited than FC27 but the differences were not statistically significant.

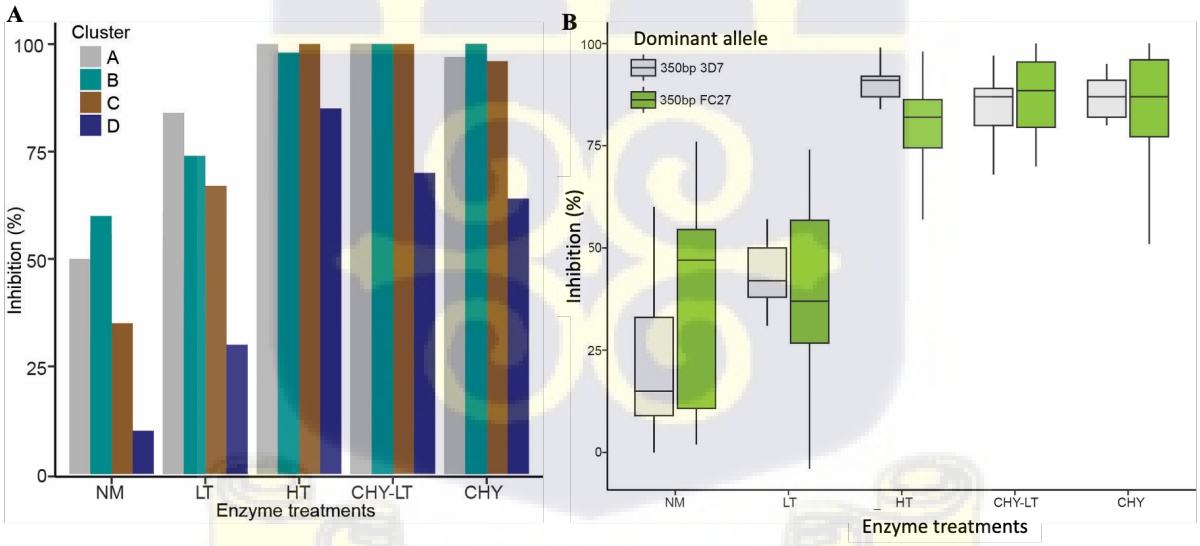


Figure 3.10: Invasion inhibition phenotypes of *P. falciparum* clinical isolates. (A) gene expression clusters, n =48 (A= 32, B = 9, C = 6, and D =1) and (B) *msp 2* dominant allele (350 bp 3D7 n = 8, 350 bp FC27 n = 23).

3.5 Discussion

This study characterised and compared the genetic diversity, ligand genes expression, and invasion phenotypes of *P. falciparum* clinical isolates across four years in The Gambia, where enhanced control interventions for over a decade have led to a significant decline in malaria transmission and clinical incidence. This overall decline in malaria transmission was reflected by lower heterozygosity in MSP2 polymorphic repeat alleles in this study except for 2022 isolates. However, world malaria reports in the recent years covered by this study indicate a slight rebound of infections in The Gambia (WHO, 2020, 2021, 2022a). Thus, an increase in polygenomic infections was observed in 2022 isolates compared to 2021, 2015, and 2016 despite the lower overall heterozygosity. Additionally, there were more infections in 2021 and 2022 with more than a single clone despite the smaller overall number of clones in circulation. This could be attributed to the slight increase in transmission reported by the recent world malaria reports. Such a phenomenon had previously been described between populations in Cameroon, where a high multiplicity of infections was observed despite the low number of circulating genotypes (Atuh et al., 2022). This observation is unexpected since The Gambia is in the pre-elimination phase of malaria and MOI is expected to be decreasing over the years. Moreover, it is generally expected that an increase in MOI should lead to increasing heterozygosity but that was not the case in this study (Fola et al., 2017; Roh et al., 2019; Touray et al., 2020). It is probable that despite the low transmission in The Gambia, diversity has reduced to a small number of different clones circulating efficiently in the population or imported in this mostly urban region of the country. However, these results should be interpreted with caution due to the small sample size, and sampling only three out of six regions in The Gambia does not reflect the entire country. Moreover, only the MSP2 gene was genotyped and the pattern could differ if diversity at other polymorphic loci such as MSP1, microsatellites, and genome-wide single nucleotide polymorphisms (SNPs) were evaluated.

Parasitemia is established in malaria following several merozoite ligand - erythrocyte receptor interactions, merozoite invasion and replication in erythrocytes (Cowman & Crabb, 2006). Erythrocyte invasion can occur through several alternative pathways, which could be determined by the expression levels, type or use of different ligand combinations (Dolan et al., 1990; Duraisingh, Maier, et al., 2003; Taylor et al., 2002). Analyses of transcript levels relative to apical membrane antigen 1 (AMA1) gene for selected invasion ligands in our clinical isolates showed significant differences in the rhoptry protein RH5, (Douglas et al., 2011; Illingworth et al., 2019), relative to EBA-181, EBA-175. RH5 is a host tropism determinant and an unavoidable ligand in invasion process, binding to Basigin (CD147), and P113 (Campeotto et al., 2020; Wanaguru et al., 2013). It is one of the leading blood-stage *P. falciparum* malaria vaccine antigens, and its higher expression relative to EBA-175 is in contrast to previous studies in Ghana and The Gambia, where dominance of EBA-175 transcripts and SA-dependent pathways were reported (Gomez-Escobar et al., 2010; Mensah-Brown et al., 2015). Unlike isolates from these previous years, most of those characterised here were SA-independent. RH5 expression correlated strongly with that of RH4 and RH2b, two ligands also important for SA-independent invasion, indicating a switch from the expression of involve in SA-dependent to those of SA-independent invasion pathways in The Gambia. Parasites may vary invasion pathways to non-EBA dependent or SA-independent pathways due to immune pressure against the EBA proteins family members as demonstrated from disruption of EBA-175 gene (Stubbs et al., 2005). Malaria has been in a decline in The Gambia, with expected reduction in population level immunity (Ceesay et al., 2008, 2010). While this should ideally reduce the pressure on EBA-175, and allow for sustained SA-dependent invasion, the reverse seen here points to immune pressure not being the only factor in pathway switch. Indeed, pathway switch to more efficient mechanisms may be conditioned by the environment as seen during *in vitro* culture adaptation (Awandare et al., 2018). High expression was also observed

with CLAG-2, a gene that is associated with the rhoptry bulb, and a member of a multigene family of rhoptry proteins, implicated in cytoadherence, infected cell permeability, and invasion (Nguitragool et al., 2011; Trenholme et al., 2000). These results together further indicate that the level of gene expression alone cannot explain the variation in patterns of invasion pathways, as these correlate with other infection matrices. For example, there was a positive correlation between RH5 expression and parasitaemia as previously reported from The Gambia (Gomez-Escobar et al., 2010), Ghana (Mensah-Brown et al., 2015), Mali (Tran et al., 2014), and Papua New Guinea (Chiu et al., 2014). RH5 is an essential invasion ligand and a potent vaccine candidate in advanced development. Its protective role against infections and disease needs to be further evaluated as malaria parasite populations evolve against intense interventions with changing epidemiology.

The increased use of SA-independent receptors and neuraminidase resistance from 2015 to 2022 in The Gambia may be because of adaptation to reduced transmission pressure. Clinical isolates of *P. falciparum* from other regions in Africa have also shown similar levels of resistance to neuraminidase treatment (Bowyer et al., 2015; Deans et al., 2007; Mensah-Brown et al., 2015). Most of these were also isolates from recent populations, in contrast to previous findings of field isolates collected over a decade ago in The Gambia, where the parasites' ability to invade neuraminidase-treated erythrocytes was much higher (Baum et al., 2003; Gomez-Escobar et al., 2010). Malaria transmission during the earlier studies was more intense in The Gambia, and the variation observed here could also be because some of the precious specimens tested were from severe malaria cases (Gomez-Escobar et al., 2010) compared to the mostly low parasitemia uncomplicated cases in the current study. There is also the possibility of variations in the activities of commercial neuraminidase enzyme batches used. Overall, the use of different invasion pathways remains heterogeneous and the molecular patterns determining

this remains complicated, needing further evaluation as parasite ligands are targeted for vaccine development.

3.6 Conclusion

Despite the variability in the mean ages of the study participants across the years (2015 and 2016 being mostly children while 2021 and 2022 were predominantly adults), this study showed that *P. falciparum* clinical isolates are mostly using sialic acid-independent invasion pathways against the sialic acid-dependent alternative. It also showed an increase in low-complexity polygenomic infections in The Gambia. With The Gambia accelerating interventions towards pre-elimination, and the possibility of new elimination tools (e.g. vaccines), continuous molecular and phenotypic surveillance of the parasite population is necessary.

3.7 Limitations

- **Temporal Sampling Bias**: The study's recruitment covered only select years—2015, 2016, 2021, and 2022—while excluding samples from the intervening years of 2017 through 2020. This selective time frame may introduce temporal sampling bias.
- The sampling population in 2015 and 2016 were mostly children while 2021 and 2022 were both children and adults and these might also introduce some level of variability in the results.

CHAPTER 4

Genetic Variations in *P. falciparum* Invasion Ligands and their Cognate Human Receptor Variants in Malaria Cases from the Gambia

4.1 Abstract

The malaria parasite *P. falciparum* invades the human erythrocytes through ligand-receptor interactions. These interactions during the erythrocytic stage of the parasite's life cycle initiate the clinical signs of malaria. The genes encoding the parasite ligands and human receptors might have experienced some evolutionary changes as a result of these interactions, and this could impact malaria incidence and outcomes in endemic areas. However, variations in parasite ligands and their corresponding receptors in the same individuals or population have received far less attention even though such studies may be useful in vaccines and drug design, refinement of control strategies as well and provide a better understanding of host-parasite interactions, disease development, and progression. To investigate this, a paired study of *P. falciparum* merozoite invasion ligands and their corresponding erythrocyte receptor genes from the same infected individual/population was carried out using the Nanopore amplicon sequencing approach. Briefly, blood samples were collected from 288 malaria-positive individuals from four health facilities in The Gambia: Brikama, Basse, Fajikunda, and Edward Francis Small Teaching Hospital (EFSTH). Genomic DNA was extracted from the whole blood and 12 *P. falciparum* genes: EBA175, EBA181, EBA140, Clag2, Clag8, Rh4, Rh5, merozoite surface protein (MSP)1, MSP6, Duffy binding-like MSP (DBLMSP), erythrocyte binding-ligand 1 (EBL-1), and surface-associated interspersed protein 4.2 (SURFIN4.2), and four human receptors: glycophorin (GP) A, GPB, GPC, and complement receptor 1 (CR1) were sequenced. Raw fastQ data were quality-checked with Nanoplot, filtered and trimmed with Porechop and NanoFilt, and then aligned with reference genomes (human: GRCh38 and *P.*

falciparum: PlasmoDB v.9.0 3D7) using Minimap2. Samtools was used to extract and sort reads, Clair3 to call variants, and VCFtools and SnpEff to filter and annotate SNPs. Samples were chosen based on minor allele frequency and missingness. Haplotype networks were computed by extracting and pseudo-phasing genotypes, converting them to fasta format, then analysing with PopArt. Custom R scripts were used for identity by state (IBS), population structure, and linkage disequilibrium (LD) analysis, which were then visualised using hierarchical heatmaps, PCA, and neighbour-joining trees.

Moderate to high levels of within-host complexity of infection across site and high inter-SNP LD were observed in the DBLMSP and SURFIN4.2 genes of *P. falciparum*, and the human glycophorin B, C, and CR1 gene. There was also a lack of spatial structure between *P. falciparum* from different sites while for the human population, individuals from Basse (Upper River Region) were relatively more distinct from the rest of the population. Analysis of the distribution of variants in receptors versus controls identified several SNPs in the CR1 receptor and a single variant in GPC associated with severe malaria. Specific host-parasite allelic combinations may determine the infection and severity of malaria. Further simultaneous understanding of these genetic variations and interaction between host and parasite may be an important approach for understanding disease pathophysiology and the development of better control interventions.

4.2 Introduction

Erythrocytes, the most abundant human blood cells serve as a decoy for bacteria and viruses to bind and eliminate them as well as a host for the malaria parasite (*Plasmodium*) (Johansson & Falk, 2021; McCullough, 2014). With *Plasmodium*, the mortality and reduced survivorship due to parasite-infected erythrocytes have resulted in selection pressure on the human genome (Kwiatkowski, 2005). This is evident by the identification of selective signatures in the genes

that code for erythrocyte surface proteins in individuals from malaria-endemic areas (Leffler et al., 2017; Thathy et al., 2005; Timmann et al., 2012). Well-established human genetic markers associated with malaria phenotypes are the sickle cell trait known to confer protection against complicated malaria (Amale et al., 2013), erythrocyte glycophorin (GP) receptors structural variants such as DUP4 (Kariuki et al., 2020; Leffler et al., 2017), *Sl2* allele of complement receptor 1 (CR1) (Thathy et al., 2005), G6PD deficiency (Mbanefo et al., 2017), FREM3, ATP2B4 (Timmann et al., 2012), as well as the Duffy negative blood group which safeguards against *P. vivax* infection (Miller et al., 1976). The selection and variations across these loci also imposed selection in the *P. falciparum* genome whereby, genes involved in erythrocyte invasion and virulence are equally variable and under evolutionary selection. Some of these genes includes EBA-175 (Verra et al., 2006), Rh2a and Rh2b (Rayner et al., 2005; Reiling et al., 2010), DBLMSP2 (Ochola et al., 2010), MSP1 (Hughes, 1992), MSP3 (Polley et al., 2007), and AMA1 (Polley & Conway, 2001). Additionally, several surface and exported genes (*Clag2, Hyp6, Hyp15, SURFIN8.2, SURFIN4.2, Clag8, Etramp10.3, and MSP7*) have signatures of selection from genome scans across many African countries (Amambua-Ngwa et al., unpublished data; Kaewthamasorn et al., 2012).

The parasite can specifically take advantage of polymorphisms in ligand genes to increase access to a variety of hosts, evade host immunity, vaccination, and antimalarial drugs (Mackinnon & Marsh, 2010; Volkman et al., 2007), making control and elimination challenging. Evasion of host immune responses can occur through allele-specific mechanisms, whereby an individual might be immune to one parasite allele that was previously encountered but not immune to a heterologous allele with no previous contact or pre-existing immunity (Lyon et al., 2008; Polley et al., 2007). During infection, rare alleles receive less immune attention and are expected to proliferate over time whereas dominant alleles are eliminated by the immunity they trigger (Conway et al., 1992). Therefore, a process of frequency-dependent

selection (balancing selection) may be responsible for maintaining polymorphisms on ligands and antigens. This diversity and several intermediate-frequency alleles/haplotypes may affect the efficacy of vaccines against merozoite antigens, due to allele-specific immunity (Genton et al., 2002; Ouattara et al., 2013).

To maintain interactions and facilitate erythrocyte invasion, merozoite ligands also need to adjust to withstand the diversity of erythrocyte receptors. Thus, dual analysis of genomics diversity in humans and parasites from endemic populations can identify genetic signatures of selection and significant evolutionary insights contributing to the success and maintenance of *P. falciparum* in humans. A better understanding of natural selection, population dynamics, and the parasite's genetic diversity will also be beneficial for the development of new or better disease control interventions. It is also possible that different variants at host receptors might have arisen in different populations because of selection by different *P. falciparum* strains, with specific receptor-ligand combinations (e.g., the erythrocyte binding proteins (EBA and EBL) and GPs) influencing parasite invasion, growth, and virulence differently. Variations in parasite ligands and their corresponding receptors in the same individuals or population have received far less attention even with the most common and well known ligands like the EBA (EBA-175, EBA-181, EBA-140, EBL-1) and Rh (Rh1, Rh2, Rh4, Rh5) protein family members and their correspnding receptors (the glycophorins, complement receptor 1, among others). It can not be over emphasised that studies like this may be useful in vaccine and drug design, as well as the development of control strategies. Additionally, continuous molecular surveillance has been recommended (WHO, 2022a) as one of the malaria elimination strategies and since the parasite is not independent of the host, a dual study on both humans and parasites will better inform public health policy. Thus, to generate baseline data on host-paraste interactions from the same study/individuals in The Gambia, the most common erythrocyte receptors, parasite ligands and surface proteins like GPA, GPB, GPC, CR1, EBA-175, EBA-181, EBA-140, EBL-1, Rh4,

Rh5, MSP1, MSP6, DBLMSP, CLAG2, CLAG8, and SURFIN4.2 were genotyped in infected individual (severe and mild malaria) with the hypothesis that specific *P.falciparum* invasion ligand and human receptor interactions may result in adaptive variants that drive infection rates.

4.3 Methods

4.3.1 Sample Size Determination

The sample size of this objective was determined using the G*Power software to ensure robust statistical analysis. G*Power is a useful tool for calculating sample sizes in genetic research because it provides appropriate power for detecting relevant genetic effects, accounts for the complexity of genetic data and models, optimises resource allocation, and supports evidence-based study design. The following parameters (effective size (f^2) = 0.2, alpha error probability (α) = 0.05, power (1- β) = 0.08, number of predictions (number of targeted genes) = 16) were used to get a minimium sample size of 200 participants for this objective.

4.3.2 Sampling population, DNA extraction, and target selection

The population sampled or sampling sites as well as inclusion and exclusion criteria were as previously described in section 3.3.3 and 3.3.4 respectively. Blood samples from a total of 288 (23 severe and 265 mild malaria) individuals were used. Genomic DNA was extracted from the packed frozen erythrocytes of each blood sample using the QIAamp® DNA Blood Mini kit (Qiagen, Cat 51104) according to the manufacturer's protocol.

Twelve *P. falciparum* invasion ligand genes (EBA175, EBL1, EBA140, EBA181, MSP6, MSP1, DBLMSP, CLAG2, CLAG8, Rh4, Rh5, and SURFIN4.2) and four human receptor genes (GPA, GPB, GPC, and CRI) were selected for genotyping by targeted amplicon sequencing. Primers targeting known polymorphic regions of each gene were selected from previous publications as indicated (appendix V, Table S3). The chromosomal locations of the 12 selected *P. falciparum* genes are shown in Figure 4.1.

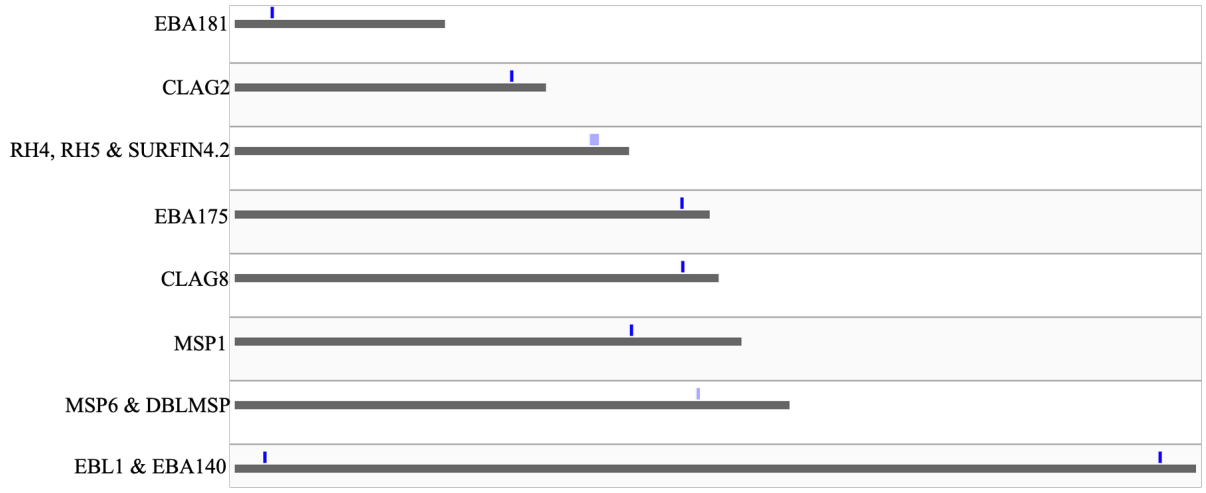


Figure 4.1: *P. falciparum* chromosomes (grey) and location of genes (blue)

4.3.3 Severe Malaria Diagnosis

Severe Malaria cases were only recruited from EFSTH which is a referral health facility in the Greater Banjul region of the Gambia. The diagnosis was done by physicians in the health facility and confirmed cases were included in the study. The WHO criteria for the definition of severe malaria were used (WHO, 2022b). These criteria included high parasitemia (parasitemia > 10 %), severe malarial anemia (haemoglobin < 7 gm/dl plus malaria positive), renal impairment (serum or plasma creatinine > 2652 µmol/L (3 mg/dL)) or blood urea > 20 mmol/L, diminished consciousness (glasgow coma score ≥ 11), acidosis, and hypoglycemia (blood glucose < 2.2 mmol/dL).

4.3.4 Amplicon generation and pooling for sequencing

All primers were first optimised separately with different polymerases using two *P. falciparum* laboratory strains 3D7 (The Netherlands) and Dd2 (Indochina). Upon successful optimisation, samples were amplified following the optimised conditions per target (appendix V, Table S3). Following PCR, successful amplification was verified for all samples and targets by gel electrophoresis (appendix V, Figure S5A and S5B) to ensure a successful amplification before pooling (appendix V, Figure S5C) for sequencing. For each sample, 3 µl of the PCR product was mixed with 2µl DNA stain and run for 45-60 minutes on a 1 % agarose gel at a voltage of

99 to confirm amplification. The amplicons were quantified using the Qubit 2.0 fluorometer with Qubit dsDNA high-sensitivity kit following the manufacturer's protocol. The molarity of the samples was then calculated and pooled in a 96-well plate. Targets with fragment sizes beyond 2 kb were pooled in a ratio of 1:3 to avoid bias during sequencing, where 1 is for targets with product sizes below 2 kb and three is for targets with product sizes above 2 kb. The Qubit dsDNA high-sensitivity kit was used to determine the final concentration of each pooled plate before library preparation.

4.3.5 Nanopore Library Preparation and Sequencing

Library preparation was done using Oxford Nanopore Technology (ONT) native barcoding expansion 96 kit (EXP-NBD196) in conjunction with the ONT ligation sequencing kit (SQK-LSK109) according to the manufacturer's protocol with slight modifications on the starting DNA concentration and the native barcode volume per sample. Summarily, at least 300 ng of input DNA (the pool of PCR amplicons from sixteen genes) was end-repaired and A-tailed using NEBNext Ultra II End-prep enzyme mix (NEB, Cat E7546) and incubated in a thermal cycler for 5 minutes at 20 ºC and 5 minutes at 65 ºC. 1 µl/sample of the end-prepped PCR amplicon mix was taken forward for barcoding and blunt end using the ONT native barcoding expansion 96 (EXP-NBD196) and 1x Blunt/TA ligase master mix (NEB, Cat M0367S) respectively. The barcoded samples were incubated for 20 minutes at 20 ºC and 10 minutes at 65 ºC. The end-prepped barcoded amplicons were pooled in a microcentrifuge tube and 480 µl of the pooled sample was purified with 1x AMPure XP beads (Beckman Coulter, Cat A63882) and eluted in 35 µl of nuclease-free water. 30 µl of the elute was ligated with NEBNext quick ligation module (NEB, Cat E6056) and adapters were added using adapter mix II expansion (AMII, ONTEXP-AM11001) and quick T4 DNA ligase (NEB, Cat E6057A) and incubated at room temperature for 10 minutes. Purification of the final library was achieved with 125 µl of short fragment buffer (ONT, Cat EXP-SFB001) twice with a magnetic stand and eluted with

15 µl of elution buffer (ONT, Cat EXP-AUX001). The library was quantified using the qubit dsDNA high-sensitivity kit. Following quantification, the library concentration, and the average fragment size of all sixteen targets were entered on the NEBioCalculator (https://nebiocalculator.neb.com/#!/ligation) to calculate the amount of library to load on the flow cell. 100-200 fmol was loaded on the R9.4.1 (FLO-MINI106) flow cell in a mix of 75 µl, constituting 37.5 µl sequencing buffer (ONT, Cat EXP-AUX001), 25.5 µl loading beads (ONT, EXP-AUX001) and 12 µl library. Sequencing was done on GridION for 72 hours with flow cell wash and library reloading every 24 hours. Samples were sequenced in three batches (96 samples per batch) and base calling was done in real-time using the high-accuracy setting (appendix V: Figure S6 and S7).

4.3.6 Data Analysis

The quality of raw fastQ files was checked using nanoplot. Reads were filtered and trimmed based on quality and length using Porechop and NanoFilt. This was followed by alignment of the human and *P. falciparum* reads to the amplicon regions of the human (GRCh38) and *P. falciparum* (PlasmoDB v.9.0 3D7) reference genomes respectively using Minimap2. Samtools was then used to extract and sort mapped reads, followed by variants calling with Clair3. After variants calling, VCFtools was used to filter and extract genotypes for further analysis, while SnpEff was used to annotate both files (*P. falciparum* and human). Using a minor allele frequency (MAF) of 0.05 and missingness of genotyped or SNPs calls at not more than 40 %, 208 out of 288 samples were retained for downstream analysis.

4.3.6.1 Haplotype Network Analysis

Genotype and allelic information were extracted using the BCFTools. The *P. falciparum* dataset was pseudo-phased by retaining the allele with the highest allelic depth at heterozygously called loci. The phased data was converted to fasta format using custom R scripts (Github, https://github.com/MPB-mrcg/Human_PF_Corelation/tree/main/Rscripts).

Using the MABL (Methodes et Algorithmes pour la Bioinformatics LIRMM), the fasta file was converted to a nexus file for haplotype generation. As the human dataset is diploid, it was used directly to generate the nexus, and the trait file containing annotation information of the samples. Both nexus and trait files were uploaded in PopArt and the haplotype network plots were generated using the TCS (Transitive Consistency Score) option.

4.3.6.2 Identity by State (IBS), Population Structure, and Linkage Disequilibrium (LD) Analysis

The unphased genotype file was analysed with customised R scripts (Github, https://github.com/MPB-mrcg/Human_PF_Corelation/tree/main/Rscripts) to generate identity-by-state (IBS) distance matrices for both *P. falciparum* and human datasets. The identity by state (IBS) metric was used to calculate the genetic relatedness between infection pairs. IBS values varied from 0 to 1, and infection pairs with an IBS > 0.5 are strongly related and equivalent to 50 % of shared alleles between the two isolates under analysis. Genetic distance matrices were generated with SNPs data of *P. falciparum* genes only, human genes only, and a combination of both *P. falciparum* and human genes using the "stamppNeisD" function in the StAMPP R package. Using the distance matrices, hierarchical heatmaps of Nei's genetic distances were generated for each dataset using the R package Pheatmap-1.0.12. Also, using genetic distance matrices, principal component analysis (PCA) and neighbour-joining (NJ) trees were generated to investigate population structure. Neighbour-joining tree estimation was done using the NJ function of the "ape" R package. The "write.function" was used to write the NJ output in parenthetic format and then read in iTOL (Interactive Tree of Life) to generate the tree for each dataset.

Linkage disequilibrium (LD) analysis was done between pairs of alleles within the same locus and across loci for human and parasite genes separately. LD was determined from the VCF files with PLINK and visualised as triangular heatmaps using the PheatMap R package.

4.4 Results

4.4.1 Characteristics of Study Participants

A total of 288 isolates were targeted for sequencing, and these were from malaria patients in The Gambia sampled from four sites Basse (69), Brikama (92), Fajikunda (69), and EFSTH (58). More males (154) were sampled than females (107), while gender was missing for 27 (9 %) individuals. Additionally, only a few (23) medically validated severe malaria cases were available for sequencing and these were only collected from EFSTH. Participants were both children and adult between 1 -70 years of age with mean ages ranging from 13.3 to 26.5 (Table 4.1). Basse was dominated by the Fulani ethnic group, while Brikama, Fajikunda and EFSTH were predominated by the Mandinka ethnic group (appendix V, Table S4).

Table 4.1: Total samples genotyped across four years

Sites	Gender			Malaria Phenotype		Mean Age ± SD
	Male (%)	Female (%)	NA (%)	Mild (%)	Severe (%)	
Basse	38 (25)	23 (22)	8	69 (26)	0	17 ± 11.7
Brikama	45 (29)	38 (36)	9	92 (35)	0	13.3 ± 8.3
Fajikunda	34 (22)	28 (26)	7	69 (26)	0	21.6 ± 12.8
EFSTH	37 (24)	18 (17)	3	35 (13)	23	26.5 ± 16.1
Total	154 (53.5)	107 (37)	27 (9.4)	265 (92)	23 (8)	

NA (not applicable): the number of participants who did not indicate their gender

SD: Standard deviation

4.4.2 Nanopore Amplicon Sequence Data Generated

Overall, a total of 34.07 million reads were generated across all samples and targets, which produced 502.54 GB of data with 28.32 GB estimated bases and 20.74 GB passed bases called. The average N50 across all three batches was 1.011kb (Table 4.2). N50 defines the contig length at which 50 % or half of the entire sequence data is present. A higher N50 is an indication of the presence of more complete and longer sequences in the data. Varying N50s were

observed across the three batches of samples, probably due to variation in amplicon proportions generated during amplicon pooling.

Table 4.2: Data generated from nanopore sequencing

Sample batch number	Number of samples per run	Run duration	Data (GB)	Total Reads (M)	Estimated bases (GB)	Estimated N50	Bases called: pass (GB)
1	96	72 hours	159.93	11.63	7.91	816 bp	5.53
2	96	71 hours 41 mins	220.39	14.47	12.18	909 bp	8.44
3	96	64 hours 12 mins	122.22	7.97	8.23	1.31 kb	6.77
Total	288	-	502.54	34.07	28.32	Average = 1.011 kb	20.74

M =million, GB = gigabyte, bp = base pairs, kb = kilo base

4.4.3 Number of Identified SNPs Per Gene

After variants calling with Clair3, VCFtools was used to filter the data using a minor allele frequency (MAF) of 0.05 (5 %), 40 % missingness per sample, and 10 % missingness per loci. SNPs that did not meet these filtration criteria were removed. A total of 208 out of 288 samples initially sequenced were retained and a total of 10 out of the 16 genes initially targeted were retained for downstream analysis after the quality control (QC) process (appendix V: Figure S6 and S7) due to fewer variants called and high level of missingness. The genes retained and their corresponding number of SNPs are shown in Table 4.3.

Table 4.3: SNPs detected in both *P. falciparum* and erythrocyte receptor genes

***P. falciparum* genes**			
Gene	**Number of SNPs**	**Synonymous SNPs**	**Non-Synonymous SNPs**
Clag2	9	3	3
SURFIN4.2	79	24	48
EBA175	4	3	0
DBLMSP	8	2	4
RH4	3	0	1
EBA140	1	1	1
Human genes			
Gene	**Number of SNPs**	**Intron variant**	**Intergenic region**
Glycophorin A	14	0	14
Glycophorin B	15	0	15
Glycophorin C	15	15	0
Complement receptor 1 (CR1)	62	62	0

4.4.4 Genetic Relatedness of Human Host and *P. falciparum* Clinical Isolates

Looking at *P. falciparum* and human genes either separately or combined, it was observed that the relationship between the isolates was predominantly independent of the geographical location and disease phenotype although each classification (*P. falciparum* only, human genes only and *P. falciparum* and human genes combined) had one cluster containing just mild malaria cases and void of the Wolof ethnic group (Figure 4.2). *P. falciparum* isolates from all four sites did not cluster by the source of samples (Figure 4.2A). When both parasite and human data were combined, population structure by source of samples was also absent (Figure 4.2C). For human SNP genotypes only, individuals from Basse clustered separately in a unique cluster void of Wolof (Figure 4.2B), thereby distinguishing Basse population and the Wolof ethnic group from the rest of the population. Two main genetic clusters were observed with the *P.*

falciparum-only SNP data (Figure 4.2A), three clusters for the human SNP data only (Figure 4.2B), and three main genetic clusters when both *P. falciparum* and human SNP datasets were combined (Figure 4.2C). The three clusters were more evident with the combined *P. falciparum* and human SNP datasets than with the human data alone. All three data sets showed overall moderate to high genetic distances between sites.

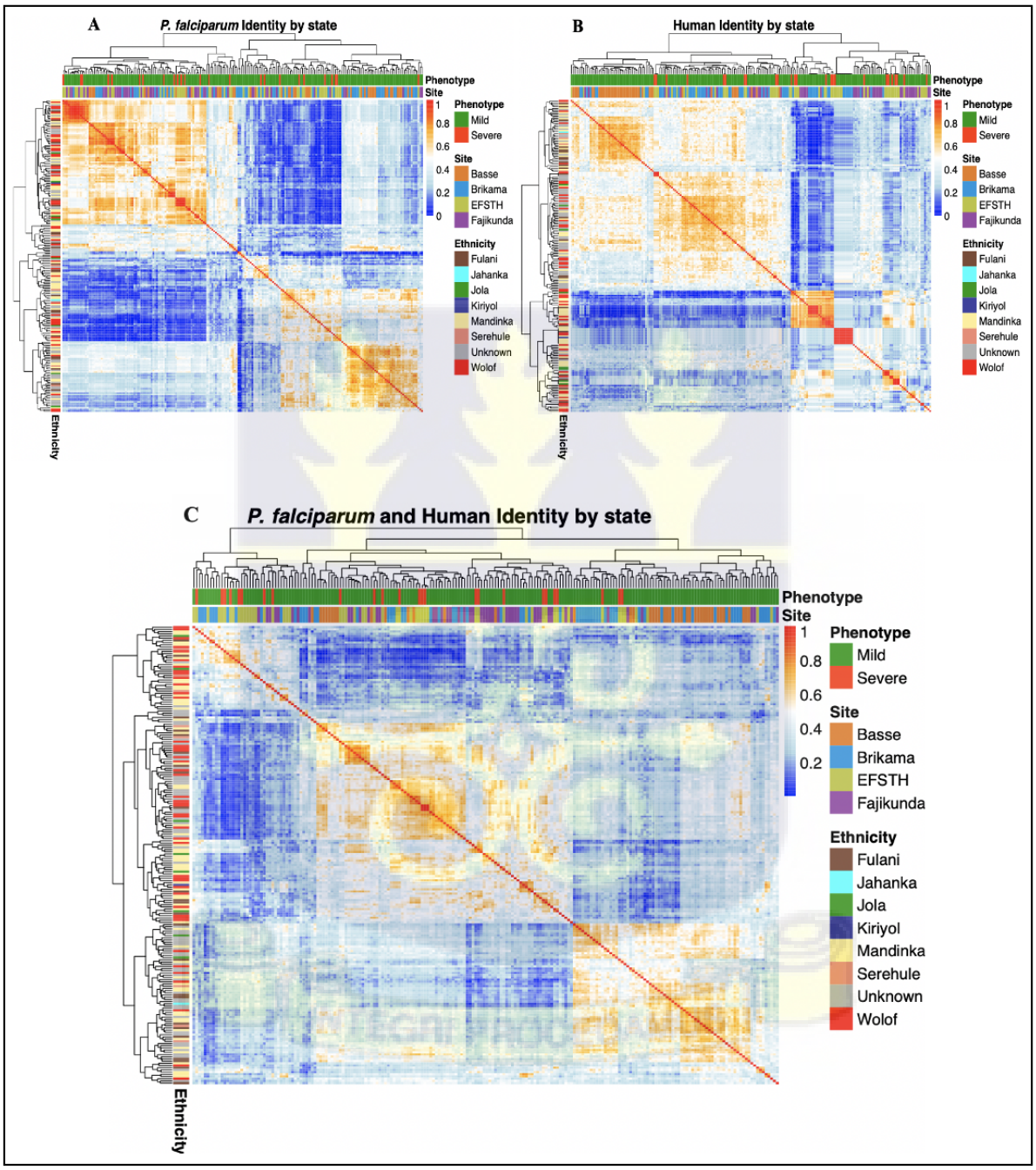


Figure 4.2: Hierarchical clustering and heatmaps based on pairwise identity-by-state (IBS) algorithm. (A) Genetic distance between samples computed from SNPs in *P. falciparum*

ligand genes only, (B) Genetic distance between samples computed from SNPs in human receptor genes only, and (C) Genetic distance between samples computed from SNPs in both *P. falciparum* ligand and human receptor genes. The phenotype and sampling site are shown as bars at the top side of the dendrograms while the ethnicity is shown on the right side of the dendrogram. The distance between individuals ranges from zero (blue) to one (red), where zero or blue indicates individuals with distinct genotypes (no similarity) while genetically identical individuals are shown in red (one).

4.4.5 Population Structure Analysis

The principal component analyses (PCA) and unrooted neighbour-joining (NJ) trees were used to further assess the population structure of the *P. falciparum* isolates and their corresponding infected human individuals. PCA was done using the pairwise genetic distance matrices generated in section 4.3.6.2. No significant population sub-structure was observed for both *P. falciparum* and humans; however, isolates from Basse were more differentiated, while those from the greater Banjul areas (including Brikama, Fajikunda, and EFSTH) showed no geographic pattern (Figure 4.3A). This lack of population structure by geographical location was evident with a neighbour-joining tree which showed five clusters but not based on the location of sampling (Figure 4.3B). For the human population, most individuals from Brikama clustered together from those recruited at EFSTH, Basse, and Fajikunda. The NJ tree of the human population further identified seven clusters that were not based on the geographical location of the individuals (Figure 4.3C and D). These findings were unchanged with a PCA (Figure 4.3E) and neighbour-joining tree (Figure 4.3F) incorporating both the *P. falciparum* and human datasets. To further explore if the observed human clusters were based on the ethnicity of the individuals from the different sampling locations, the distribution of ethnicities of the sampled population per site was computed and it was observed that Basse was dominated

by the Fulani and Mandinka ethnic groups while the other sites were dominated by the Mandinka and Wolof ethnic groups (appendix V: Table S4).

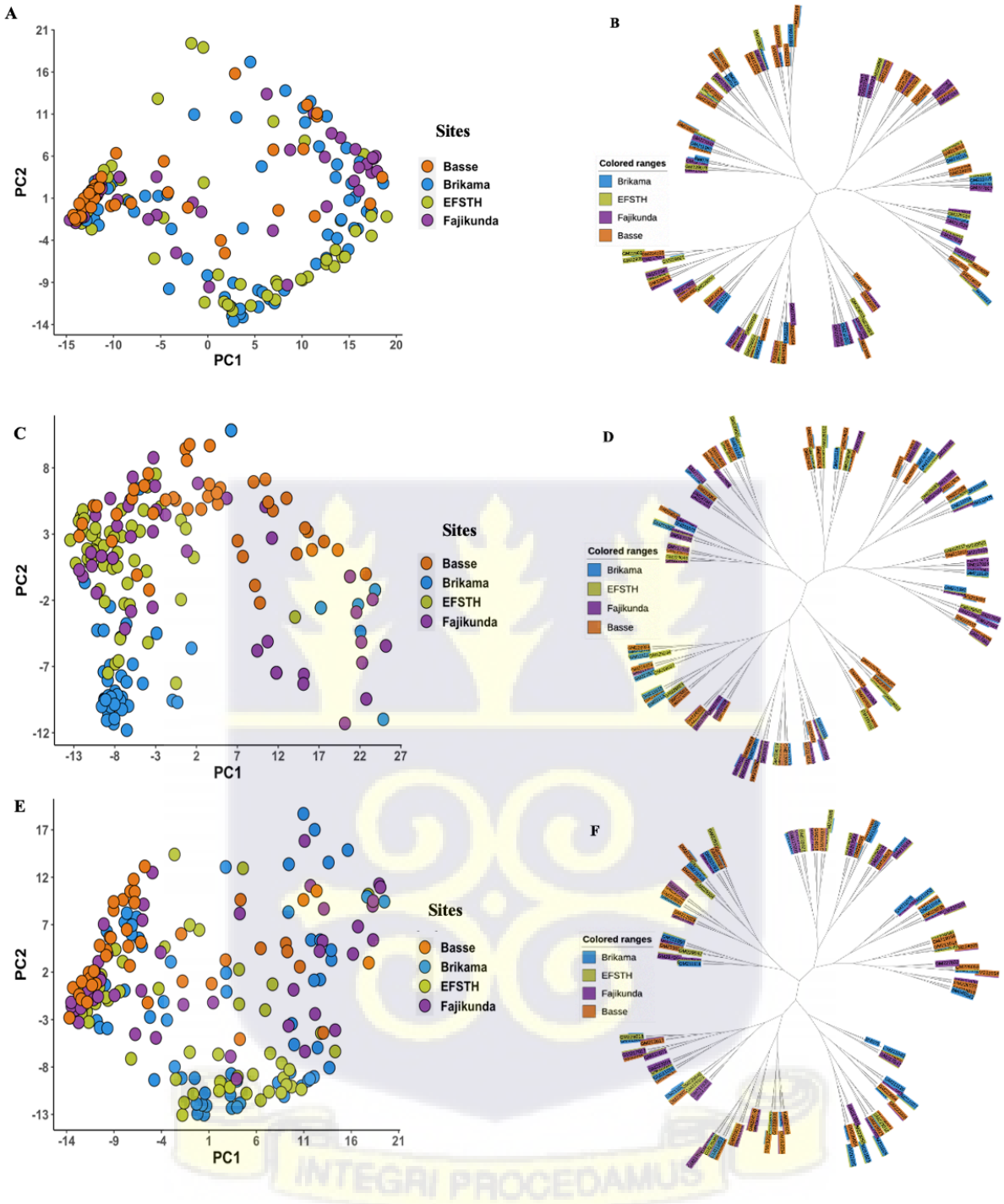


Figure 4.3: Population structure analysis of both *P. falciparum* and human genes based on sampling sites. (A) Principal component analysis (PCA) plot of *P. falciparum* genes, (B) Unrooted neighbour-joining (NJ) tree for *P. falciparum* coloured, (C) PCA plot of human

receptors, and (D) Unrooted NJ tree for human receptors, (E) PCA plot of human receptors and *P. falciparum* genes, and (F) Unrooted NJ tree of human receptors and *P. falciparum* genes.

To further understand these results, the population structure analysis was carried out per gene for both *P. falciparum* (Figure 4.4) and human populations (Figure 4.5). No significant population structure was observed for each gene. However, 4 clusters of *P. falciparum* isolates were observed with SURFIN4.2, one of which was predominated by isolates from Brikama and EFSTH (Figure 4.4A). For human populations, the glycophorin C (GPC) gene clustered individuals into 2 main distinct clusters, one containing all Brikama individuals, mixed with individuals from other sites while the second cluster mainly contained individuals from Basse, Fajikunda, and EFSTH (Figure 4.5C). A few SNPs from EFSTH were also distant from the rest of the populations probably because isolate from EFSTH were mostly from severe cases. Further statistical analysis of GPC variants against disease severity identified 2 SNPs at position 126693612 (A-G) and position 126693617 (A-T) with high chi-square values (13.4 and 15.7 respectively). Previously reported SNP (A-C) at position 126693612 was different from the one (A-G) identified here. Data from the Gambian genome variation project downloaded from the Ensembl database showed that the two identified SNPs were not previously reported in The Gambia population from microarray genotyped and whole genome sequence data.

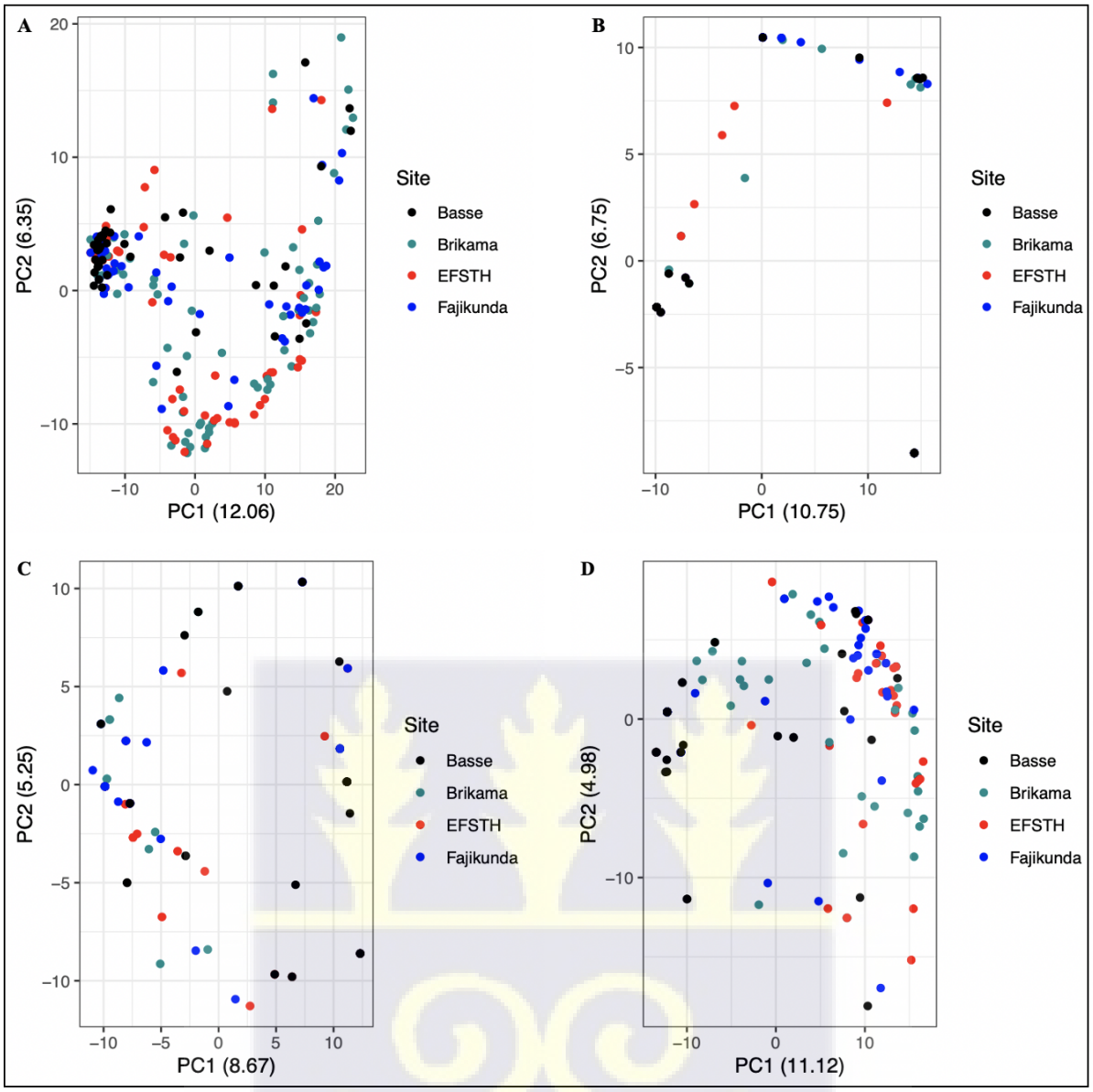


Figure 4.4: Principal component analysis (PCA) of some *P. falciparum* genes. (A) SURFIN4.2, (B) DBLMSP, (C), EBA175, and (D) CLAG2

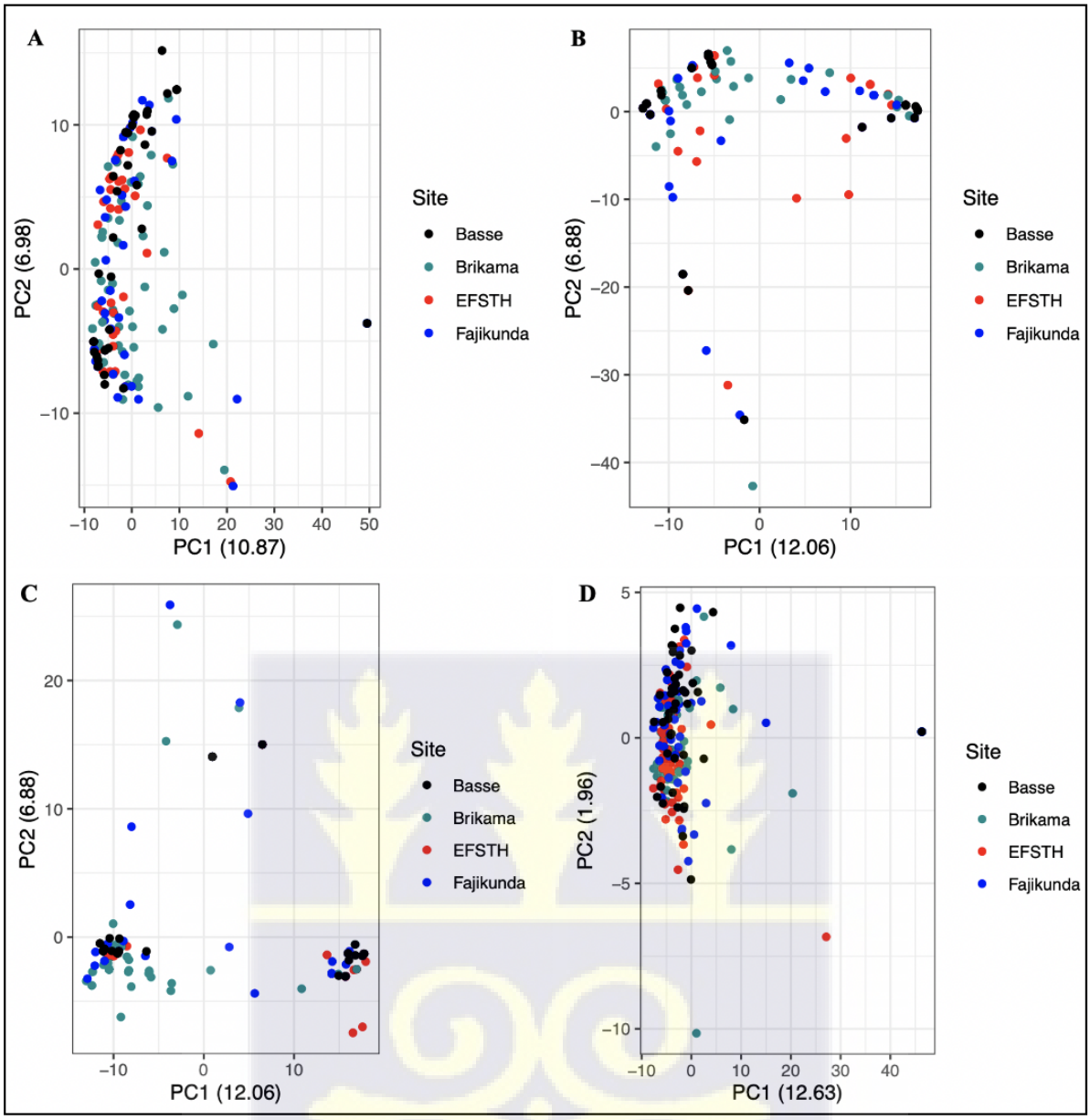


Figure 4.5: Principal component analysis (PCA) of human receptors. (A) Glycophorin A, (B) Glycophorin B, (C) Glycophorin C, and (D) Complement receptor 1 (CR1)

4.4.6 Complexity and Population diversity of *P. falciparum* infection

The likelihood that any two randomly selected infections have distinct alleles at a given locus is expressed by Wright's inbreeding coefficient (*Fws)*. *Fws* was computed using the formula: $Fws = 1 - \frac{Hw}{Hs}$ where Hs is the heterozygosity of the population sampled, and Hw is the heterozygosity of the infection across all loci. Mixed infections defined by *Fws* values less than

0.95 were observed in the parasite SNP data of all four sites sampled. The complexity of infection was similar across all sites with majority of the samples having *Fws* values below 0.5, indicating mostly complex infections across all sites (Figure 4.6).

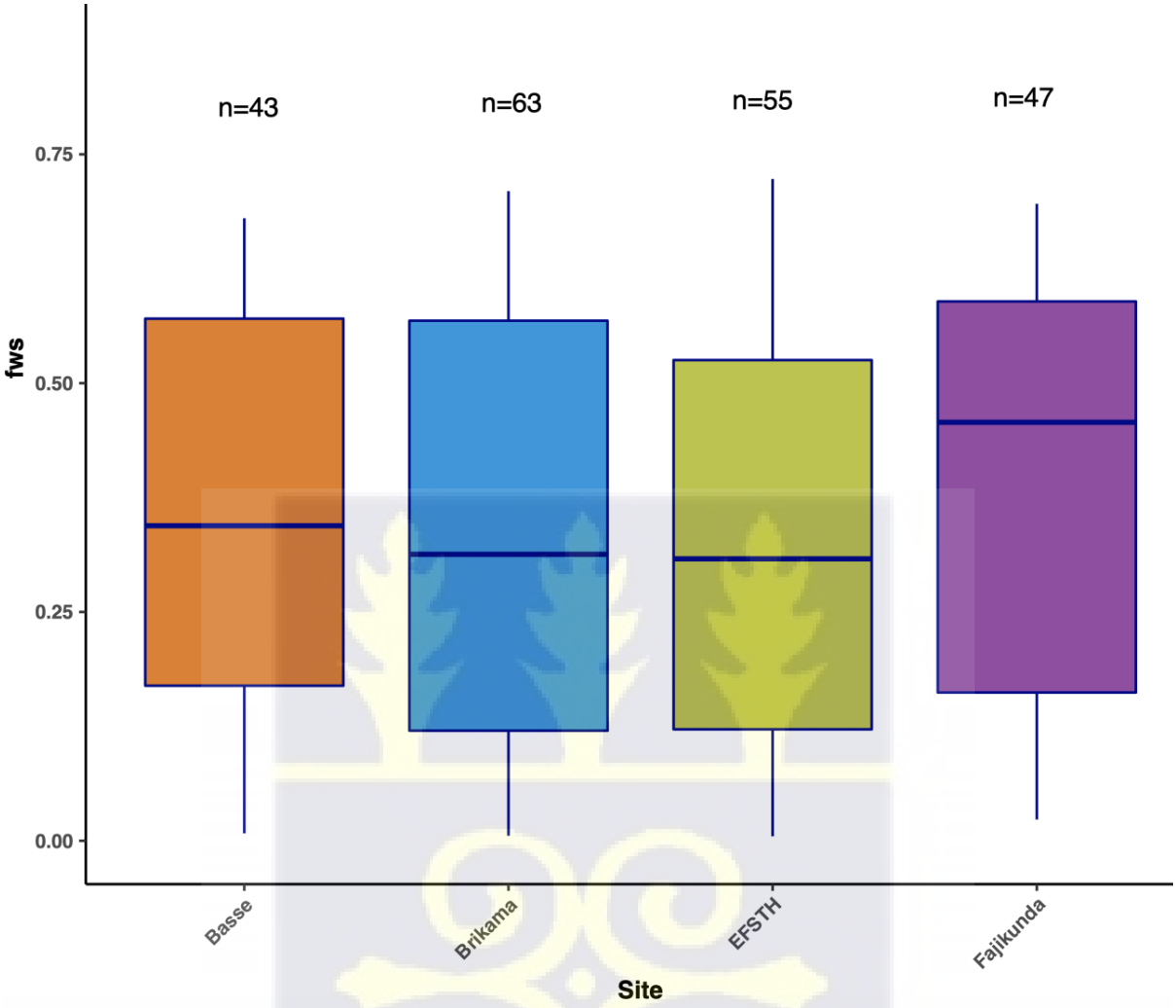


Figure 4.6: Distribution of Wright's inbreeding coefficient (*Fws*) estimates in samples from each of the four study population.

4.4.7 Linkage Disequilibrium Analysis between human and *P. falciparum* Genes

Significant and varying multilocus linkage disequilibrium (LD) patterns were observed between SNPs on various *P. falciparum* and human genes as shown by the LD heatmaps (Figure 4.7). For *P. falciparum* LD (Figure 4.7A), SNPs in chromosome 10 (the DBLMSP gene) (block 3) had the highest r^2 values, indicating a strong association between the alleles at

this loci. High r^2 values were also observed between most of the SNPs in chromosome 4 (SURFIN4.2 gene) (block 1 and 2) compared to other genes like the EBA-175 on chromosome 7, EBA-140 on chromosome 13, and Clag2 on chromosome 2 with only one significant SNP each. For LD within human genes (Figure 4.7B), high r^2 values were observed between all SNPs in the glycophorin B gene (block 1), some SNPs in the CR1 gene (block 2), and very few SNPs in glycophorin C (block 3) of the human genome. All genes not reported here did not have significant LD values.

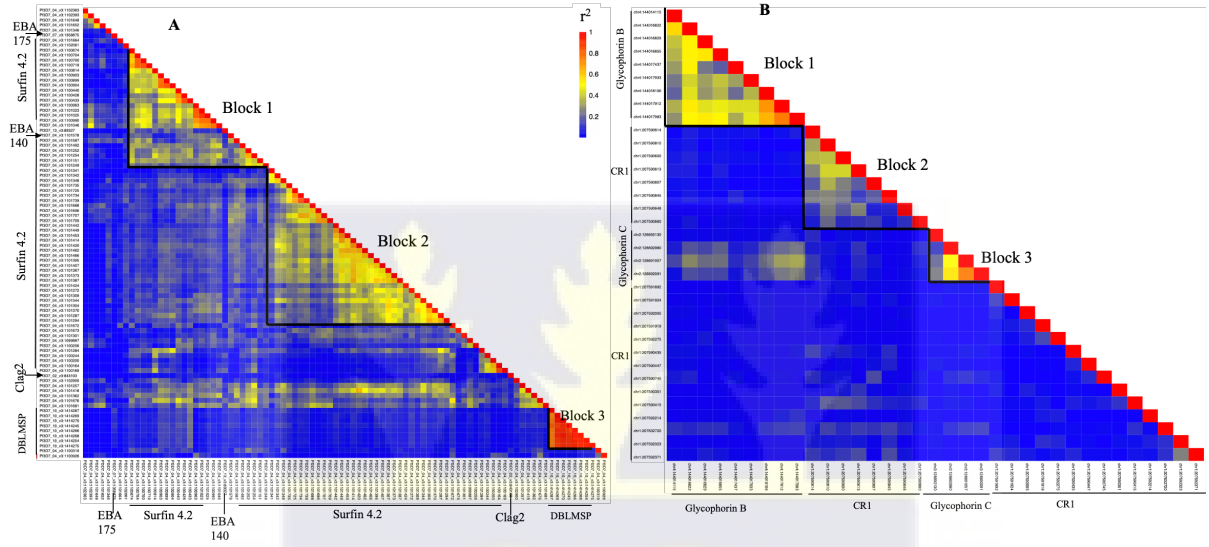


Figure 4.7: Linkage disequilibrium (LD) heatmap between pairs of SNPs generated using the Pheatmap R package and measured by r^2. (A) LD between SNPs on different *P. falciparum* ligand genes, (B) LD between SNPs on different human receptors. r^2 = allelic values' squared correlation at two loci.

4.4.8 Tajima's D Test for Neutrality

The Tajima's D test was used to determine deviation from neutrality. Tajima's D values were computed using PLINK. Tajima's D is a neutrality test that takes into consideration the discrepancy between the expected Watterson's population nucleotide diversity (theta or θ) under neutrality from the number of segregating sites (S) and the average pairwise nucleotide diversity between sequences (π) (Tajima, 1989). Positive Tajima's D values were obtained

from two *P. falciparum* SURFIN4.2, and DBLMSP. For the human genes, CR1, glycophorin B and C had positive Tajima's D values (Table 4.4). The positive Tajima's D values indicate that these genes are under balancing selection.

Table 4.4: Computed Tajima's D values of both *P. falciparum* and human genes

***P. falciparum* Genes**			**Human Genes**		
Gene	**Number of SNPs**	**Tajima D**	**Gene**	**Number of SNPs**	**Tajima D**
SURFIN4.2	79	5.66162	**CR1**	22	4.61145
DBLMSP	8	3.1882	**Glycophorin C**	4	1.37677
			Glycophorin B	9	2.24185

4.4.9 Haplotype network analysis of *P. falciparum* and human Genes

4.4.9.1 *P. falciparum* Haplotype Networks

The Haplotype network allows for a clearer observation of intra-species nucleotide sequence variation due to variation between the sequences and the potential of recombination events. Networks were constructed for each gene using the "population analysis with reticulate trees (PopART) package (Leigh and Bryant, 2015). The *Clag2* gene had 6 main haplotype blocks shared between sites and the rest were unique to a single site. Three haplotypes (Hap_1, Hap_2, and Hap_5) were shared between all four sites, 2 (Hap_4 and Hap_6) were shared between Brikama, Fajikunda, and EFSTH, while 1 (Hap_3) between Fajikunda, Basse, and EFSTH. Hap_1 was the dominant block (n = 63), with the majority (n = 29) originating from Fajikunda (Figure 4.8A). The MSP6, and DBLMSP genes were more diverse with mostly singletons unique to single sites (Figure 4.8B and D). This high level of diversity was expected because these are highly polymorphic genes. SURFIN4.2 had the highest number of haplotype blocks shared between sites (Figure 4.8C) which was expected because it equally had the highest number of SNPs among all the *P. falciparum* genes.

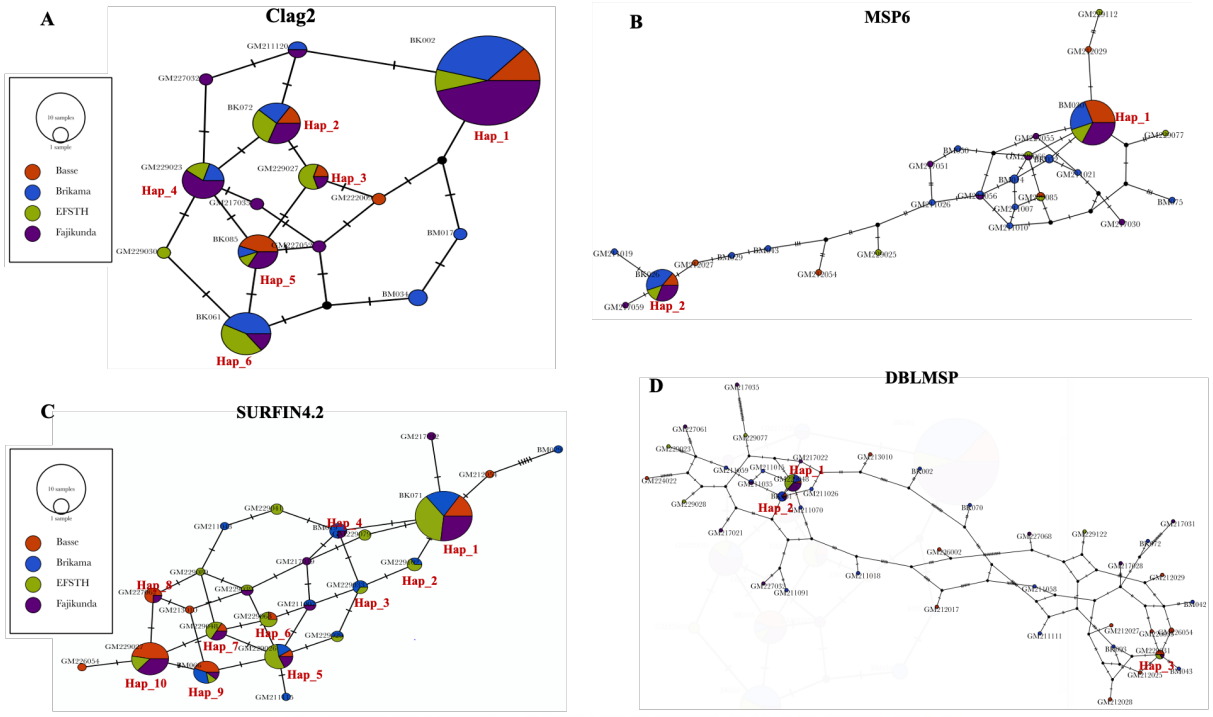


Figure 4.8: Haplotype networks of *P. falciparum* genes. The colours represent the geographical distribution of the haplotypes, the circle size of each haplotype represents its frequency and the hatch marks between cycles indicate the number of mutations per haplotype.

4.4.9.2 Erythrocyte Receptors Haplotype Networks

The same approach used to generate haplotype networks for *P. falciparum* was used to generate networks for the human genes. The network plot for glycophorin A (GPA) revealed 11 main blocks, 5 (Hap_2, Hap_7, Hap_8, Hap_9, and Hap_11). Hap-2 was the most dominant block (n = 40), with the majority (n = 14) originating from EFSTH (Figure 4.9A). Glycophorin B (GPB) on the other hand had fewer (7) haplotypes shared between sites but many singletons unique to various locations. Three (Hap_1, Hap_4, and Hap_7) of the 7 main haplotypes were shared between all four sampling sites and Hap_1 was the most dominant (n = 50) and included 19 individuals from Brikama (Figure 4.9B). GPC had six haplotype blocks shared between all four sites and it was hihly polymorphic with relatively more SNPs per haplotype compared to other human genes (Figure 4.9C). Complement receptor 1 (CR1) gene was the most diverse with the highest number of shared haplotypes between sites (Figure 4.9D).

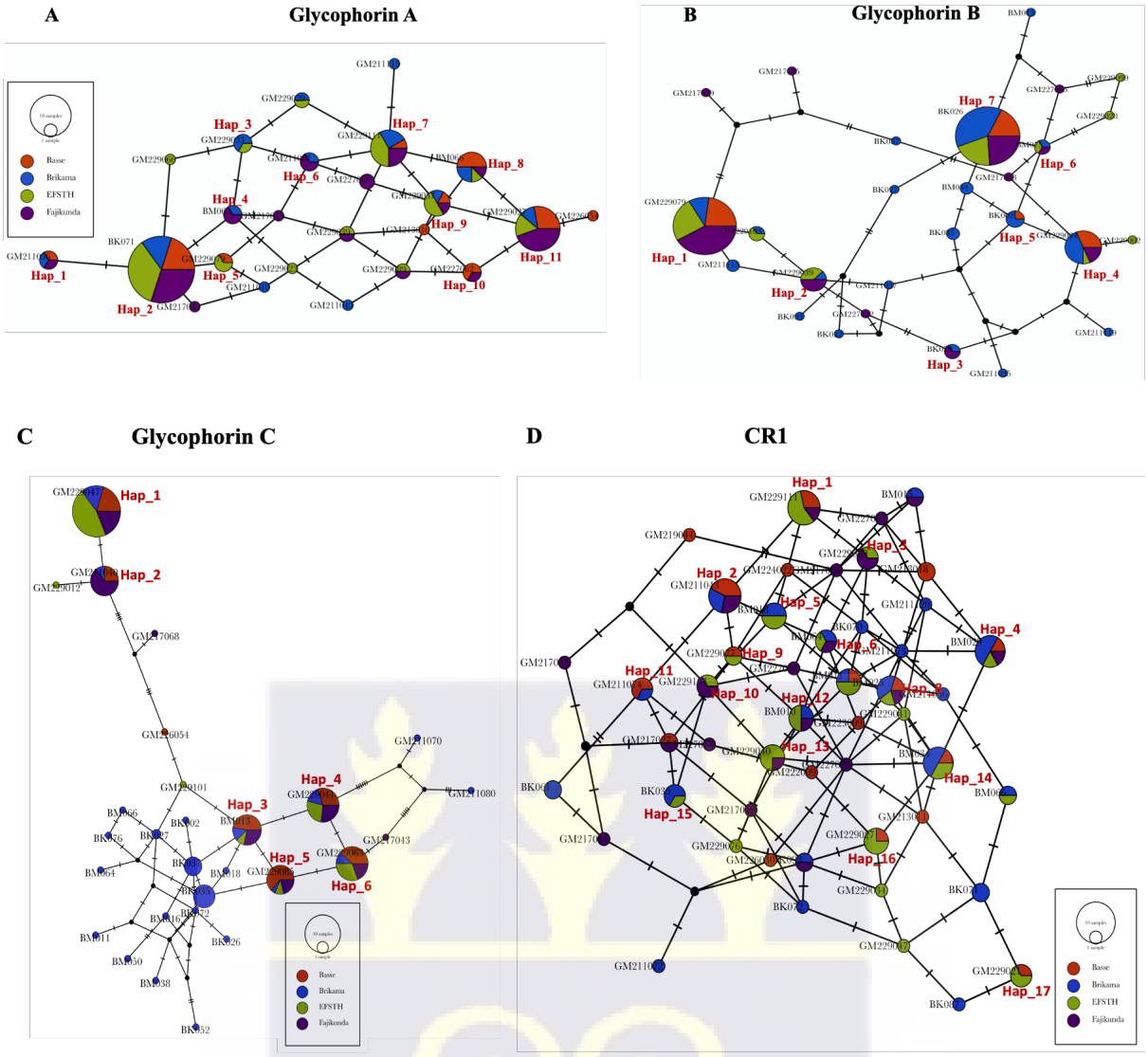


Figure 4.9: Haplotype networks of erythrocyte receptors. The colours represent the geographical distribution of the haplotypes, the circle size of each haplotype represents its frequency and the hatch marks between cycles indicate the number of mutations per haplotype.

4.4.10 Relationship Between Identified Loci and Severe Malaria

Univariable logistic regression models for case (severe malaria)-control (mild malaria) was used to determine if the identified loci of each gene (both human and *P.falciparum*) were associated with severe malaria. Seven SNPs in CR1 were strongly associated with malaria susceptibility (P values ≤ 0.05). A single SNP in SURFIN 4.2 also had a significant association with malaria susceptibility. Few SNPS in both CR1 and GPC were marginally significant (P values = 0.069), while the majority SNPs across all genes were insignificant (Figure 4.10). Out

of the seven significant SNPs identified in CR1, four (labeled) have previously been identified but their relationship with malaria severity have not been established while three (unlabeled) have not been reported before. Also, the significant SNP in the SURFIN 4.2 have not been previously reported and this is the first time a SNP in the parasite gene is found to be associated with malaria severity. A table with all SNPs per gene indicating their levels of significance has been provided in appendix V, Table S5. It is important to highlight that acute kidney injury, renal impairment, and cerebral malaria were the most prevalent forms of severe malaria, and these conditions largely accounted for the observed associations.

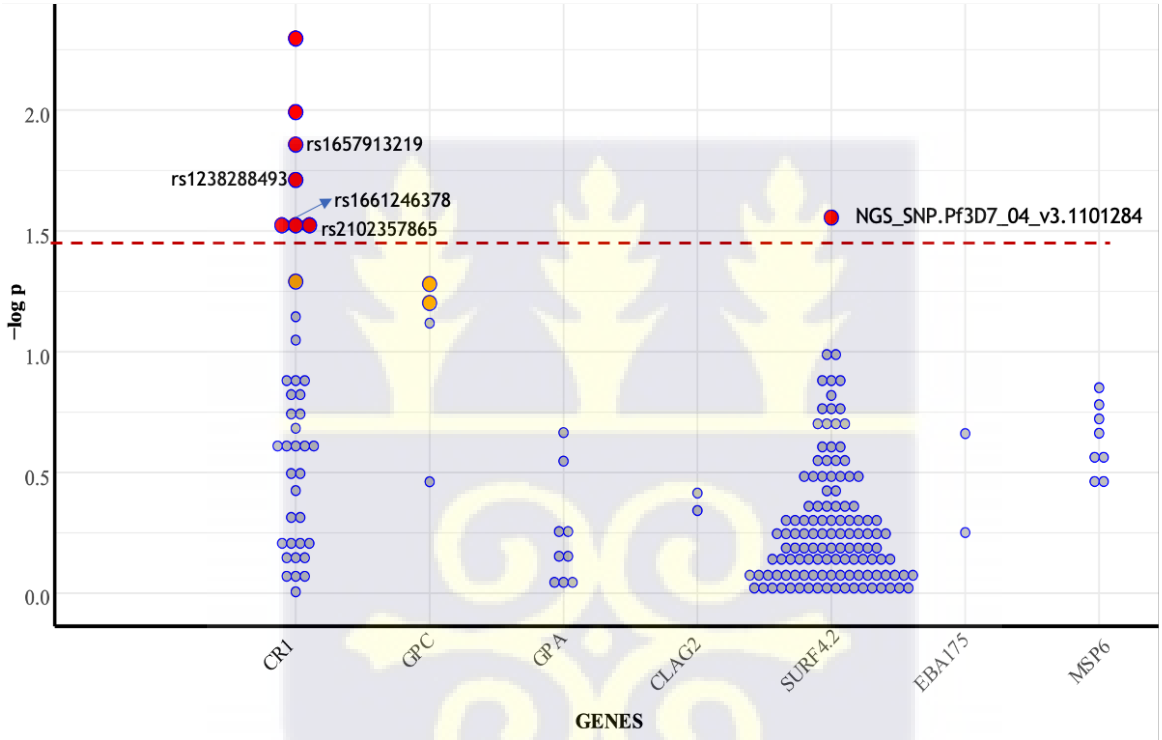


Figure 4.10: Relationship between identified loci and severe malaria. Red dots represent significant SNPs ($P \leq 0.05$), orange dots = marginally significant SNPs ($P = 0.069$) , and blue dots = nonsignificant SNPs.

4.5 Discussion

Receptor-ligand interactions during the erythrocytic cycle of *P. falciparum* is one of the factors influencing the evolution of malaria parasites and human hosts. Thus, malaria parasite and host genome architectures have been shaped by the coevolution of genes functioning at the interphase of invasion. While variations in the parasite ligands determine the parasite's ability to properly invade the erythrocyte and adapt to environmental factors and pressure from control strategies, variations in the erythrocyte receptors determine the ability of the host to disrupt the invasion process and prevent the parasite from establishing in the host. Therefore, analysing dual genetic variations in both the host and the parasite in the same individual from malaria endemic populations will increase our understanding of their role in the erythrocytic cycle, disease pathogenesis, and genetic mechanisms of protection from infection or severe disease. To this effect, a targeted genomic approach by nanopore was used to simultaneously describe genetic variants in parasite invasion ligands and human receptors in individuals from The Gambia.

The population structure analysis of *P. falciparum* markers did not identify distinct clusters based on the geographical location of the isolates/individuals, indicating weak or absence of population structure (similar genetic makeup). As seen across most African countries, parasite populations have been mixing due to migration and a strong substructure was not evident even between the most distant populations such as isolates from Basse compared to the rest of the coastal populations, separated by over 350 km. Recent genetic surveillance studies revealed a significant flow of parasites across The Gambia (Amambua-Ngwa et al., 2019). Malaria prevalence in the Basse area is higher and has been seeding infections to the western urban areas through human migration (Mwesigwa et al., 2017). For *P. falciparum*, improved transportation across The Gambia makes it easier for people to travel and conduct businesses throughout the country leading to the admixture of *P. falciparum* strains. Contrary to the

observation for parasites, variants in the erythrocyte receptors clustered the Brikama and Basse population into a distinct sub-groups from other sites based on relatedness determined by identity by state. As seen in the demographic characteristics of study participants, the Basse population is dominated by the Fulani ethnic group and inbreeding is higher. The Fulani have been shown by previous Gambian genotyping studies to be distinct from other local populations (Jallow et al., 2009). The weak or no population structure from the coastal sites may be due to the cosmopolitan nature of the urban centres population migration.

Two SNPs in the human glycophorin C gene were significantly associated with severe malaria despite the small sample size. The GPC gene commonly known as the Gerbich blood group contains several mutations previously reported to be associated with malaria severity in endemic areas (Jaskiewicz et al., 2018). Recently, most studies on the human glycophorin genes have been on the glycophorin ABE (GPABE) locus on chromosome four due to the presence of large structural variants in this locus (Leffler et al., 2017; MalariaGEN, 2015). As a result, GPC on chromosome two has received less attention. These results suggest that GPC may also be co-evolving with its corresponding parasite ligand EB-140. Unfortunately, the interaction between GPC and EBA variants in severe and mild infections could not be further analysed as only one significant SNP was retained for *P. falciparum* ligand EBA-140 following quality filtration of the data. Further analysis of this ligand-receptor pair (GPC/EBA-140) could clarify the interactions and dual variants.

Strong LD was observed between variants in some *P. falciparum* markers (DBLMSP and Srufin 4.2) and also in human receptors (glycophorin B, Glycophorin C, and CR1), which might indicate high frequencies of a limited number of haplotypes which could be maintained by selection. SURFIN 4.2 is known to interact with the GLURP (Glutamate-Rich Protein) and RON4 (a rhoptry neck protein) to mediate contact of the merozoite and the erythrocyte during merozoite invasion. It has been shown by extended haplotype homozygosity to be under

positive selection in multiple populations. Therefore, the high level of polymorphisms observed in this gene might be an adaptation mechanism for immune evasion and other host and environmental factors.

Positive Tajima's D values were observed for three human and two *P. falciparum* genes, indicating that these genes are under balancing selection. For the parasite, this included SURFIN 4.2 and DBLMSP, consistent with previous findings on these genes (Amambua-Ngwa et al., 2012; Ochola-Oyier et al., 2019; Van Tyne et al., 2011). These are members of erythrocyte invasion and immune protein families that have previously been identified to be under balancing selection (Amambua-Ngwa et al., 2012; Van Tyne et al., 2011). The high level of polymorphisms in the DBLMSP and SURFIN 4.2 genes observed in this study might be due to the already existing selective pressure from the immune mechanism, given their direct implication in invasion pathways and malaria pathology (Quintana et al., 2018; X. Zhu et al., 2017). As antibodies against this locus inhibit invasion, SURFIN 4.2 warrants further consideration for development as components of a future multivalent vaccine against malaria.

The identification of SNPs in the CR1 gene associated with malaria susceptability was a true relationship because CR1 polymorphisms have previously been reported to play a role in malaria complications either through clearance of immune complexes (Cosio et al., 1990) or promotion of rosette formation (Rowe et al., 1997). In The Gambia, a previous study has shown that a SNP (S1 2 McC^{b}) in the CR1 gene underlying the knops blood group is positively selected in the population, though not associated with severe malaria (Zimmerman et al., 2003). Thus, the identification of significant SNPs in the CR1 gene associated with severe malaria in this study is indicative of possible other genetic variants that are associated with malaria in The Gambia and other African populations that need further functional evaluation.

4.6 Conclusion

The parasite as well as the human population have been mixing continuously, with the arms race between parasite ligands and receptors resulting in significant diversity in their haplotypes in *P. falciparum* and the human population respectively. High inter-SNP LD and multiple haplotypes at both *P. falciparum* ligand genes and the human glycophorin B, C, and CR1 genes were observed. Also, some of the variants in these ligands and receptors affect the outcome of disease, and this could be selected for malaria susceptibility. CR1 and glycophorin C seem to have important roles in malaria outcomes that warrant further investigation. These findings suggest that studying genetic variations in both host and parasite simultaneously could improve understanding of malaria parasite biology and inform the development of better control interventions.

4.7 Limitations and Recommendations

There are several genes involved in the interaction between malaria parasites and erythrocytes. Only a small number were selected here, mostly driven by convenience and a focus on known ligand-receptor interaction at the invasion stage of the erythrocytic cycle. Hence, understanding host-parasite interactions needs a wider dual genomic study. Based on available resources, targeted sequencing on ONT was the technology of choice, although this has not been largely used for malaria beyond targeting drug resistance loci. Thus, a significant amount of development was required without adequate time to validate the approach for each target.

The case-control analysis was limited by the few number of severe cases, collected only from a single health facility. A wider understanding of these interactions will benefit from an African-wide study and severe, mild, and asymptomatic malaria parasite infections. Given resources, whole genome sequencing approach to target both the human host and parasite in the same individual might provide more insight in host-parasite interactions.

CHAPTER 5

Distribution of Glycophorin B Receptor Variants Among Malaria Cases in The Gambia

5.1 Abstract

Evolutionary events on the human glycophorin B (GPB) locus, have resulted in the deletion of different regions of the gene (GPB deletion 1 to deletion 8), numbered according to their prevalence in the human population. In West Africa, deletion 1 and deletion 2 (DEL 1 and DEL 2) are the most common GPB deletions with allele frequencies ranging between 5 - 15 %. Given the function of the glycophorins as receptors for *P. falciparum* invasion, it is important to determine the distribution of these variants in malaria-endemic areas and their effect on malaria severity. Although different assays have been developed to genotype these variants, their distribution among malaria-infected individuals in The Gambia is still unknown. Using restriction fragment length polymorphism (RFLP) and multiplex PCR approaches, GPB DEL 1 and DEL 2 were genotyped in 522 malaria-positive individuals (severe and mild) from three (Upper River, Western, and Greater Banjul) regions in The Gambia. All individuals with any or both deletions were also genotyped for heamoglobin SS (HbSS), the sickle cell variant. Seven ethnic groups were examined and descriptive analysis was used to determine the allele frequencies of DEL 1 and DEL 2 variants while the Fisher's Exact test was used to test for association with malaria clinical phenotypes.

The overall allele frequencies for GPB DEL 1 and DEL 2 were 9.92 % and 1.7 % respectively. For DEL 1, the Fulani and Balanta ethnic groups had higher allele frequencies (14.8 % and 14.3 % respectively), while the Serehule and Mandinka had higher frequencies of DEL 2 (2.5 % and 2.4 % respectively). Additionally, higher proportions of both DEL 1 (6.3 % heterozygous and 2.1 % homozygous) and DEL 2 (0.4 % heterozygous and 0.6 % homozygous)

were observed in participants from Basse in the Upper River Region, which is the region with the highest malaria burden in the country, highly populated by the Fulani ethnic group. HbSS was found to be co-evolving in population with GPB deletions, suggesting a join role of the two variants in malaria. Though limited by power due to small number of severe malaria samples, an association analysis with the deletion and severe malaria found no association with malaria severity. In conclusion, there is a high frequency of GPB DEL 1 in The Gambia and this is most prevalent in the Fulani ethnic group. Both DEL 1 and DEL 2 are more prevalent in areas of high malaria prevalence but the deletions showed no association with severe malaria probably due to a limited number of severe cases.

5.2 Introduction

Malaria exert a strong selection pressure on the human genome, resulting in several protective single nucleotide polymorphisms (SNPs) and structural variants. Prominent amongst these SNPs are hemoglobinopathies such as haemoglobin S (HbS) and C (HbC) (Amale et al., 2013), glucose-6-phosphate dehydrogenase (G6PD) deficiency (Hedrick, 2011), thalassemias, structural variants like the recent Dantu variant (Kariuki et al., 2020; Leffler et al., 201; MalariaGEN, 2015), *FREM3* and A*TP2B4* genes that have also been associated with malaria susceptibility (Mbanefo et al., 2017; Timmann et al., 2012)). Recently, a structural variant in the human GPA-B-E gene cluster, known as duplication four (DUP4) and commonly found in people with the Dantu blood group antigen has been linked to a 74 % and 43 % decrease in the risk of developing severe malaria among homozygous and heterozygous individuals respectively (Leffler et al., 2017). Besides DUP4, several other structural variants have been identified in the glycophorin locus due to the high recombination and similarity between homologues in this gene family. One such variant is the deletion of either different sections or the whole GPB gene. The frequency of these polymorphisms differs between populations probably driven by the environment, ancestry, and the history of malaria. In West Africa, GPB

deletion 1 (DEL 1) and deletion 2 (DEL 2) are the most common among a total of eight deletions (DEL 1 to DEL 8) identified in the glycophorin B gene so far (Leffler et al., 2017). Given that glycophorins are abundant on the surface of RBCs and their role in *P. falciparum* invasion, they are critical for the erythrocytic cycle of the parasites, which is mostly responsible for the disease. Some variants like DUP4 have been associated with malaria disease and shown to protect against severe malaria. Unlike DUP4, the other structural variants, including the deletions, are yet to be characterised and their association with malaria established. With DEL 1 and DEL 2 being the most common in West Africa, it is important to determine their distribution and association with malaria in different West African countries. This can be done using recently developed PCR-based genotyping approaches targeting these two most common deletions (Amuzu et al., 2021; Lane et al., 2020). Compared to genome sequencing approaches, these PCR-based approaches are cheaper and do not require complex bioinformatics pipelines for data analysis. Using PCR-based genotyping approaches, GPB DEL 1 and DEL 2 variants were genotyped and their distributions among malaria-positive individuals determined across three regions in The Gambia.

5.3 Methods

5.3.1 Inclusion/Exclusion Criteria

The inclusion criteria for this project was people with fever (axillary temperature ≥ 37.5 °C) or history of fever in the last 48 hours who consulted at any of the health facilities between September to December (malaria season) of each sampling year and consented to participate. were screened for malaria infection. Those who tested positive for *P. falciparum* malaria by immunochromatic rapid diagnostic testing and reported not to have taken antimalarial drugs during the preceding three days were also invited to participate on the study.

5.3.2 Sample Size Determination

The sample size for this objective was determined using Cochran's sample size formula (Cochran, 1977) bellow taking into consideration the estimated glycophorin B deletion prevalence of 15 % in West Africa and a margin of error of 3 % in this research. The Z-value for a level of confidence of 95 % (1.96) was used to determine an initial sample size and then it was corrected for a finite number of 2.8 million people which is the population of The Gambia. This calculation yielded a final sample size of 544.

$n = \frac{Z^2.\ p.\ (1-p)}{E^2}$ Where: Z is the Z-value for a 95 % confidence level (1.96), p is the estimated prevalence of glycophorin B deletions in West Africa (0.15), and E is the margin of error (0.03).

5.3.3 Sampling Population and Diagnosis of Severe Malaria

Samplling population and sites were the same as described in section 3.3.3. Although the calculated sample size was 544 participants, a variety of practical challenges, such as participant availability and health facility capacity led to the recruitment of 522 participants. However, this slight reduction in sample size still maintains adequate statistical power and does not significantly compromise the study's validity. Out of these 522 participants sampled, 23 were severe malaria cases all from EFSTH while 499 were mild or uncomplicated malaria cases, giving an overall percentage of 95.6% uncomplicated cases and 4.4 % severe malaria cases. Recruitment was done in three years; 2020 (Brikama and Fajikunda), 2021 (Basse, Brikama, and Fajikunda), and 2022 (Basse, Brikama, Fajikunda, and EFSTH), since this was the study/project period. 2 ml of whole blood was collected from each RDT-positive and consented individual. Thin film smears were prepared, stained with 10 % Giemsa, and observed under the light microscope to further confirm malaria positivity while part of the whole blood was used for DNA extraction and GPB deletions genotyping.

Severe malaria was diagnosed as previously described in section 4.3.3 and a total of 23 severe malaria cases were confirmed and genotyped for the two glycophorin B variants (deletion 1 and 2).

5.3.4 DNA Extraction and GPB deletion 1 (DEL 1) and deletion 2 (DEL 2) Genotyping

Genomic DNA was extracted from whole blood using the QIAamp DNA blood mini kit (Qiagen, Cat 51306). Genotyping of GPB DEL 1 and DEL 2 variants was done using PCR-restriction fragment length polymorphism (RFLP) assays as previously described (Amuzu et al., 2021). Using variant-specific primers, the PCRs were performed in a final volume of 10 µl, following the conditions indicated in appendix V, table S6. Gel electrophoresis was used to confirm amplification and the restriction enzymes AciI (NEB, Cat 101229-196) and BsrBI (NEB, Cat 101228-596) were used to digest the PCR products for DEL 1 and DEL 2 respectively as shown in appendix V, table S7. The digested products were then separated by size using 1 % agarose gel prepared with 1x TBE buffer and stained with SYBR Safe DNA gel stain (Invitrogen, Cat S33102). Summarily, 2 µl of 6x DNA gel loading dye (Thermo Fisher Scientific, Cat R0611) was added to 5 µl of each digested PCR product and loaded onto the gel. 1x TBE buffer was used to run the gels at 100V for 60 minutes using a gel system (BIO-RAD USA, Cat 7000082). The gels were visualised using the Gel Doc™ XR+ imaging system (BIO-RAD) and the fragment sizes were determined using a 1kb plus molecular marker (NEB, Cat N0469S) (appendix V, Figure S8)

5.3.5 Confirmation of Glycophorin B Deletion 1 and 2 by Multiplex PCR

A multiplex PCR protocol (Lane et al., 2020) for DEL 1 and DEL 2 with no enzyme digestion step was used to double confirm all positive samples genotyped using the PCR-RFLP method described above. Briefly, a total of 72 samples either positive for glycophorin B DEL 1 or DEL 2 or both were genotyped using the primers and conditions in appendix V, table S8. The PCR products were then separated by size using 2 % agarose prepared with 1x TBE buffer and

stained with SYBR Safe DNA gel stain (Invitrogen, Cat S33102). The gels were visualised using the Gel Doc™ XR+ imaging system and the fragment sizes were determined using a 1 kb plus molecular marker (NEB, N0469S) (appendix V, Figure S9). The results from the two protocols were compared and positive samples from both protocols were considered positive for either one or both deletions.

5.3.6 HbSS (Sickle cell) Genotyping

Each individual with any of the deletions was genotyped for sickle cell since this is the most common genetic dosorder associated with more severe clinical outcomes and mortality compared to other genetic disorders. Briefly, specific primers were used to amplify the region of interest as previously described (Caroca & De Lima, 2016). The PCR was carried out in a final volume of 10 µl, following the conditions indicated in appendix V, table S9. The PCR product was digested with Ddel enzyme (R0175L: New England Biolabs). Following digestion, the product was separated with QIAxcel (Qiagen) and fragment sizes were determined using QIAxcel ScreenGel Software (appendix V, Figure S10).

5.3.7 Data analysis

All gel images were scored as heterozygous, homozygous, or wild type for both DEL 1 and DEL 2 on an Excel sheet using the banding pattern of positive and negative controls of each of the alleles. Scoring was done by two people to minimise errors. The R statistical package version 4.2.0 was used for the entire analysis. The Wilcoxon rank-sum test was used to compute or assess statistical differences and qualitative variables are presented as proportions. The allele frequencies of glycophorin B DEL 1 and DEL 2 were computed using the formular $\frac{n(aa)+\frac{1}{2}n(Aa)}{AA+Aa+aa}$, where n = the number of samples, aa = homozygous allele, Aa = heterozygous allele, and AA = wild type. The Fisher's Exact test was used for association analysis between severe and mild malaria. P values <0.05 were considered significant.

5.4 Results

5.4.1 Characteristic of Study Population

Sampling was done across three years (2020-2022) from four Health facilities (Basse, Brikama, Fajikunda, and Edward Francis Small Teaching Hospital (EFSTH). A total of 522 samples were collected: 499 (95.6 %) from mild malaria and 23 (4.4 %) from severe malaria individuals. Most of the participants, 218 (41.8 %) were from Basse in the Upper River region, followed by Brikama 169 (32.4 %), and Fajikunda 79 (15%) in the Western region, and then EFSTH 56 (10.7 %) in the Greater Banjul region. Samples from EFSTH were only collected in 2022 while in Fajikunda, sampling was done in 2021 and 2022 (Figure 5.1A).

Participants were from seven ethnic groups in the Gambia: Fulani, Mandinka, Wolof, Serehule, Jola, Balanta, and Johanka. The most represented ethnic group was the Mandinka 181 (34.8 %), followed by the Fulani 137 (26.3 %), while the least represented were the Jahanka 7 (1.3 %), and the Balanta 6 (1.2 %) respectively (Figure 5.1B). Majority of the Fulani were from the Basse area in the Upper River region. The gender distribution of study participants was slightly skewed as 54.6 % (285) were male while 43.5 % (227) were female and the gender of 1.9 % (10) individuals was missing. This difference between male and female participants was statistically significant (P = 0.009) (Figure 5.1C). Both children and adults were targeted and the participants' ages ranged from 1 to 70 years with a mean plus/minus standard deviation of 20.6 ± 13.5SD and a median of 18 (Figure 5.1D).

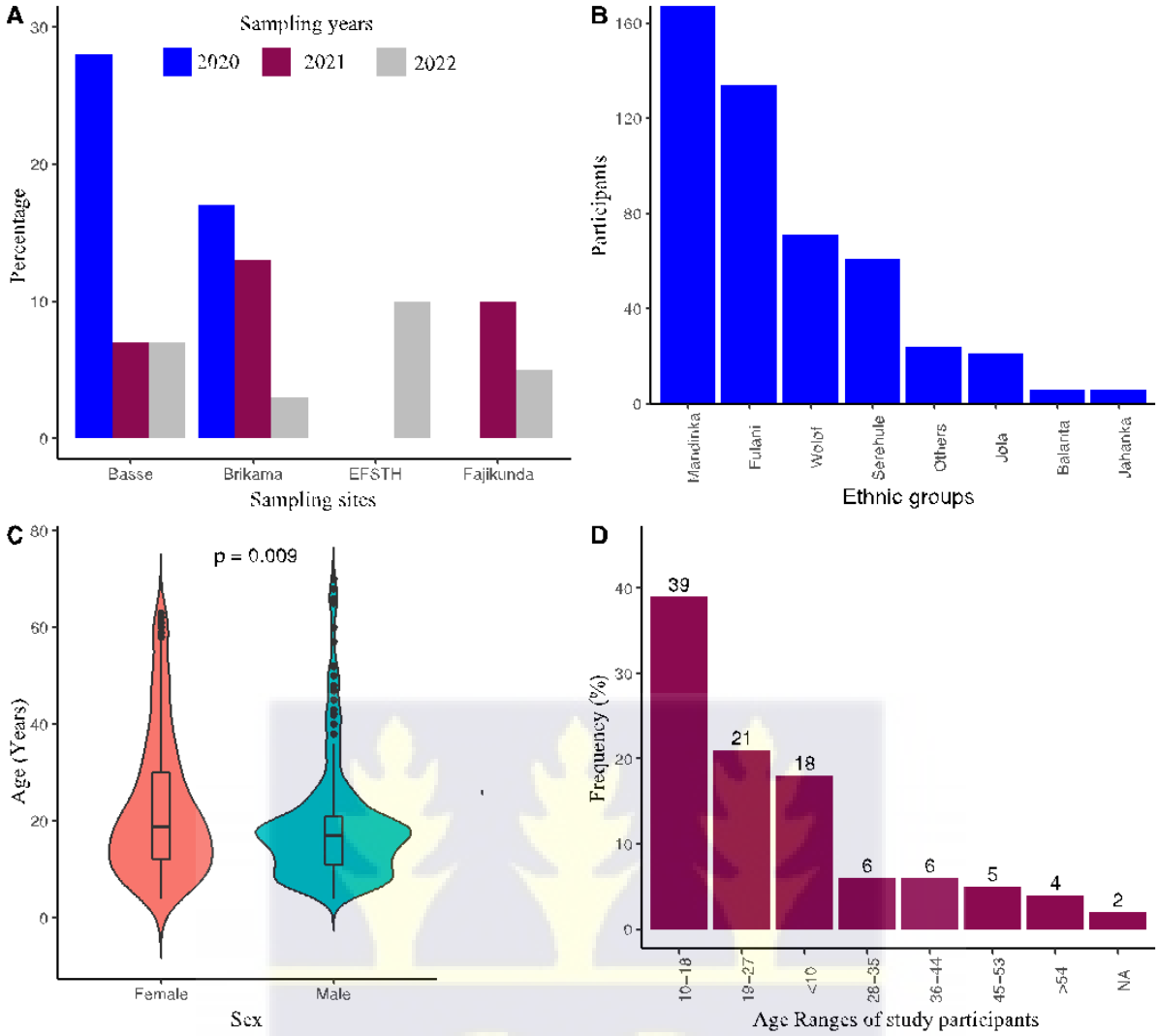


Figure 5.1: Demographic distribution of study participants. A) distribution by sampling years B) distribution by ethnicity, C) distribution by gender, and D) distribution by age groups.

5.4.2 Prevalence of Glycophorin B Deletion 1 and 2 in The Gambia

Out of the 522 samples collected, 490 (93.9 %) were successfully genotyped for deletion 2 (DEL 2) while only 383 (73.4 %) were successfully genotyped for deletion 1 (DEL 1). Out of these successfully genotyped samples, 44 (12 %) were heterozygous, 16 (4 %) homozygous, and 323 (84 %) were wild-type for DEL 1 (Figure 5.2A). DEL 2 had 7 (1 %) heterozygous, 5 (1 %) homozygous, and 478 (98 %) wild type (Figure 5.2B). Hence, the prevalence of DEL 1 in The Gambia was higher than that of DEL 2.

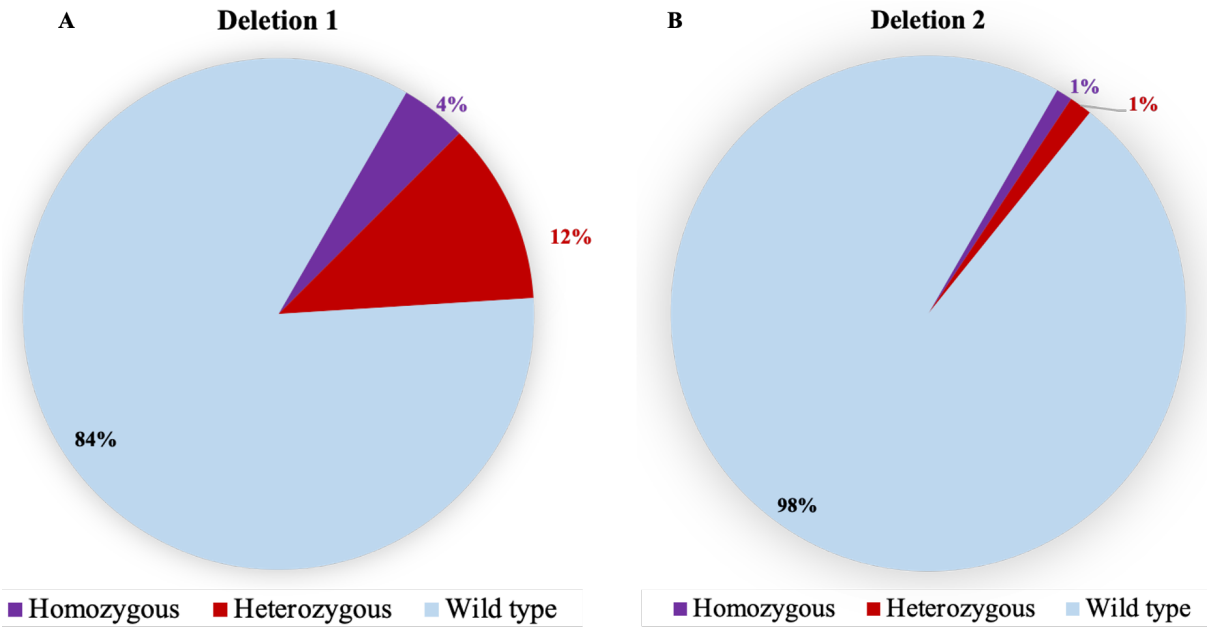


Figure 5.2: Prevalence of glycophorin B deletions in The Gambia, A) Deletion 1, B) Deletion 2

5.4.3 Allele Frequencies and Distribution of Glycophorin B Deletions by Ethnicity

The observed allele frequency of DEL 1 in The Gambia was 9.92 %. Based on ethnicity, the Fulani had the highest frequency followed by the Balanta and the Jola (14.8 %, 14.3 %, and 11.5 % respectively), while the Jahanka had the least (0 %). The heterozygous 44/383 (11.5 %) allele was more prevalent in the population than the homozygous 16 (4.2 %) (Table 5.1) but the difference was not statistically significant (P = 0.32). Unlike DEL 1, the overall allele frequency of DEL 2 was 1.7 % and this variant was only found among the Serehule (2.5 %), the Mandinka (2.4 %), and the Fulani (2.2 %) (Table 5.2). Heterozygous individuals (7), were slightly more than homozygous individuals (5) but the difference was not statistically significant (P = 0.95). One person was heterozygous for both deletions. Overall, the differences in allele frequencies across ethnic groups in both DEL 1 and DEL 2 were not statistically signicicant (P-values = 0.23 and 0.3 respectively).

Table 5.1: Distribution of DEL 1 by ethnicity

Ethnicity	Homozygous	Heterozygous	Wild Type	Total	Allele frequency (%)
Wolof	0	9	45	54	**8.3**
Fulani	9	16	90	115	**14.8**
Jahanka	0	0	4	4	**0**
Jola	0	3	10	13	**11.5**
Mandinka	5	8	100	113	**8**
Serehule	0	8	50	58	**6.9**
Balanta	1	0	6	7	**14.3**
Others	1	0	18	19	**5.3**
Total	**16**	**44**	**323**	**383**	**9.92**

Table 5.2: Distribution of DEL 2 by ethnicity

Ethnicity	Homozygous	Heterozygous	Wild Type	Total	Allele frequency (%)
Wolof	0	0	71	71	**0.0**
Fulani	2	2	130	134	**2.2**
Jahanka	0	0	6	6	**0**
Jola	0	0	21	21	**0**
Mandinka	2	4	161	167	**2.4**
Serehule	1	1	59	61	**2.5**
Balanta	0	0	6	6	**0**
Others	0	0	24	24	**0**
Total	**5**	**7**	**478**	**490**	**1.7**

5.4.4 Distribution of HbSS Among Individuals with Glycophorin B Deletions

A total of 72 individuals with either DEL 1 and/or DEL 2 were genotyped for HbSS. Seven individuals with DEL 1 were positive for HbSS while only 1 person with DEL 2 was positive for HbSS, giving a total of 8 (11.11 %) HbSS-positive individuals among the 72 genotyped. Notably, no HbAS-positive individual was found among the 72 people genotyped (Table 5.3).

Table 5.3: Distribution of HbSS among individuals with glycophorin B deletions

	DEL 1 (%)	DEL 2 (%)	Total
HbSS	7 (11.3)	1 (8.3)	8
HbAS	0	0	0
HbAA	53 (88.3)	11 (91.7)	64
Total	60	12	72

5.4.5 Distribution of Glycophorin B deletions by Gender and Study Site

Analysis of glycophorin B deletion 1 and 2 by gender indicated a higher prevalence of both DEL 1 and DEL 2 homozygous and heterozygous in males than in females (Figure 5.3A). Moreover, Basse had the highest prevalence of both DEL 1 heterozygous (6.3 %) and homozygous (2.1 %) followed by Brikama, EFSTH, and then Fajikunda. Likewise, the prevalence of DEL 2 homozygous (0.6 %) was also higher in Basse compared to Fajikunda (0.4 %) and EFSTH (0.3 %) (Figure 5.3B). The high prevalence of both deletions in Basse was due to the high Fulani population in that area.

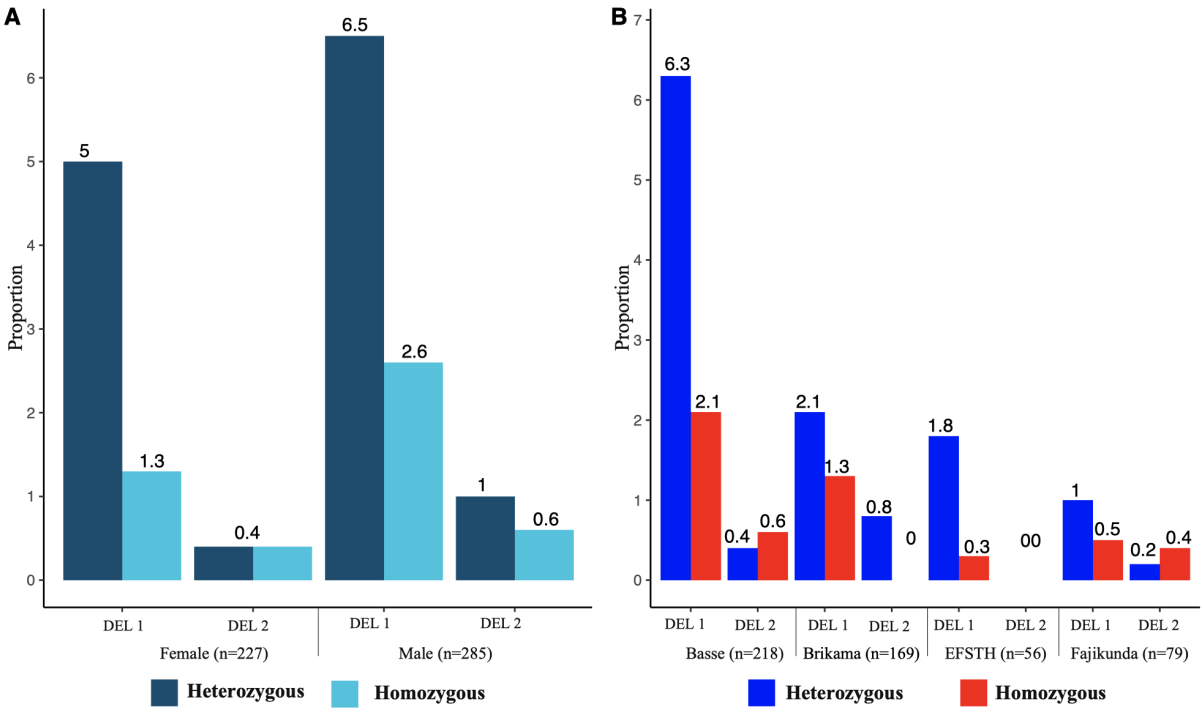


Figure 5.3: GPB DEL 1 and DEL 2 distribution (A) gender and (B) sampling sites

5.4.6 Characteristics of Severe Malaria Cases

Almost all individuals with severe malaria were adults with an average age of 26 years. Among these, severe malaria with acute kidney injury (AKI), renal impairment, and cerebral malaria (CM) were the most prevalent forms of severe malaria. Additionally, male cases (17/23) were more than female (6/23) (Table 5.4).

Table 5.4: Biochemical and haematological characteristics of severe malaria individuals

Sample ID	Age	Sex	CM	AKI	Urea	Creatinin	Anemia	Hb
GM229085	18	Male	No	Yes	44.6	568.1	No	10.6
GM229089	18	Male	No	Yes	24.4	219.3	No	10
GM229090	18	Male	No	Yes	44.6	309.3	No	8
GM229111	40	Male	No	NA	NA	NA	NA	NA
GM229022	17	Male	No	Yes	44.6	786.7	No	11.4
GM229024	18	Male	No	Yes	40.3	440.6	No	13.4
GM229025	27	Female	No	NA	NA	NA	No	10.5
GM229030	42	Female	No	Yes	44.2	293.3	No	12.6
GM229040	18	Male	Yes	No	9.21	86.2	No	13.3
GM229055	40	Male	Yes	No	NA	NA	No	11.6
GM229061	15	Female	Yes	Yes	32.2	488	Yes	5.8
GM229062	60	Female	Yes	Yes	21.89	208.8	No	11.6
GM229076	17	Male	No	Yes	36.9	388.4	No	11

GM229088	20	Male	No	Yes	44.6	724.9	No	11
GM229074	31	Male	No	No	10.6	126.9	No	10
GM229077	20	Male	No	No	21.1	194.2	No	9.2
GM229082	19	Male	No	No	6	62.2	No	10.5
GM229102	25	Male	No	Yes	NA	1138	No	13
GM229107	22	Male	No	No	6.3	61.8	No	7.9
GM229118	20	Male	Yes	No	6.5	107.3	No	NA
GM229104	15	Male	Yes	No	8.7	73	No	12
GM229086	15	Female	Yes	Yes	15.9	600	Yes	7
GM229099	61	Female	No	Yes	NA	951.1	No	9.8
Mean	**25.9**		**Yes = 7**	**Yes = 13**	**25.7**	**391.41**	**Yes = 2**	**10.49**

Hb = Heamoglobin, AKI = Acute Kidney Injury, CM = Cerebral Malaria, NA, Not applicable

5.4.7 Distribution of Glycophorin B Deletions by Malaria Phenotypes

Severe malaria cases were only recruited from EFSTH. Out of 23 severe malaria cases recruited, 22 (95.7 %) were successfully genotyped for DEL 1 while 19 were successfully genotyped for DEL 2. Among those genotyped, 1 (4.55 %) was homozygous for DEL 1 while 2 (9.1 %) were heterozygous. There was no DEL 2 in the severe cases genotyped (Table 5.5).

Table 5.5: DEL 1 and DEL 2 distribution among severe malaria individuals

Variant	Homozygous	Heterozygous	Wild type
Deletion 1 (%)	1 (4.55)	2 (9.1)	19 (86.4)
Deletion 2 (%)	0	0	19 (100)

5.4.8 Association Analyses with Severe Malaria

Given the small number of severe malaria cases compared to mild, and that only 3 DEL 1 cases were identified out of the 23 severe cases genotyped, it was challenging to do an association analysis with severe malaria. However, an attempt was made by randomly selecting a matched set of mild cases, followed by Fisher's exact test. No association was found between DEL 1 and severe malaria (P = 1.00).

5.5 Discussion

Genotyping of the two human glycophorin B deletions (DEL 1 and DEL 2) in four areas of The Gambia indicated higher frequencies of DEL 1 compared to DEL 2. This was expected because DEL 1 was classified as the most abundant variant in West Africa. Also, the frequency of DEL 1 and DEL 2 were comparable to earlier estimates of frequency ranging between 1-15 % based on analysis of 1000 Genomes and HapMap datasets (Leffler et al., 2017, MalariaGEN, 2015). The observed frequencies of these deletions varied across the seven ethnic groups represented with the Fulani having the highest frequency of DEL 1, followed by the Balanta, and the Jola, also consistent with analysis of Leffler et al., 2017 which reported that the Fulani harbors the highest frequency of DEL 1 in West Africa, compared to other ethnic groups. The Fulani ethnic group has been reported to have inherent resistance to malaria (Fall et al., 2023; Quin et al., 2017; Troye-Blomberg et al., 2020). Therefore, glycophorin B DEL 1 may be contributing to mechanisms of resistance to malaria in the Fulani population. Functional studies will be needed to further determine the mechanisms and test this hypothesis. The data showed a higher prevalence of both DEL 1 and DEL 2 in males compared to females. Males were more represented in the study than females and these were clinically malaria individuals. With males risking more malaria probably due to behavior, the data could therefore be biased by recruitment from the clinic, although the allele frequencies aligned with previous reports (Benegal, 2005; Lefèvre et al., 2010). Less bias estimates could be from community surveys and this will enable determining the effect of glycophorin B deletion on asymptomatic parasite carriage.

Basse, situated in the Upper River region of The Gambia harboured the highest frequencies of DEL 1 and DEL 2. Malaria transmission persists in Basse compared to other regions of the country (Mwesigwa et al., 2017). However, Basse population is dominated by the Fulani people, who have developed genetic resistance to malaria. High malaria burden has been a major source of selective pressure on the human genome and could have historically contributed to the current pattern in the distribution of the deletion in different regions and ethnicities (Kwiatkowski, 2005).

The sickle cell genotyping results indicated that HbSS might be linked to GPB deletions and these genotypes (HbSS and GPB deletions) might be circulating together in the population. Since HbAS offers protection against malaria and HbAS individuals are usually healthy, they are less likely to visit the hospitals hence, the absence of HbAS in the results. Studies have shown that sickle cell have greatly shaped malaria prevalence due to the sickling nature of the RBCs (Bunn, 2013; Mwaiswelo et al., 2020). Other studies have also shown that glycophorin diversity has been shaped by the presence of malaria parasites (Kariuki et al., 2020; Louzada et al., 2020). These divergent studies therefore suggest that both the sickle cell gene and the glycophorin gene might have been evolving independently. However, the results of this study show that the evolutionary mechanism might be linked together. Therefore, further studies are needed to explore the mechanism by which HbSS evolution is linked to GPB deletions in malaria-endemic areas.

No association was found between deletion 1 and severe malaria. This is the first study with an attempt to test for the association of glycophorin B deletions with severe malaria. Although this attempt was limited by the sample size, continuous monitoring of these mutations is required especially in studies with larger sample sizes both in and out of The Gambia.

5.6 Conclusion

Overall, the allele frequency of GPB deletions in The Gambia was 9.92 % and 1.7 % for DEL 1 and DEL 2 respectively. The frequency of GPB DEL 1 was highest in the Fulani, who mostly live in high malaria burden eastern regions of the country (The Upper River Region). HbSS and the GPB deletion variants might be linked or co-evolving in population, suggesting a join protective effect against malaria. An association between any the deletions and severe malaria was limited by the small number of severe malaria cases. Therefore, future studies should include more severe malaria cases to enable association analysis between the deletions and severe malaria.

5.7 Limitations

- The primary limitation of this study was the clinical-based recruitment which fails to capture asymptomatics and only a small number of severe malaria cases were recruited. This limited the power in understanding the association between the deletion and malaria phenotype spectrum (symptomatic mild and severe, asymptomatics and control individuals who are malaria negative).

- Sampling was opportunistic and therefore not uniform across study sites and years, and only three out of six regions of the country were represented. Hence, results should be interpreted with caution.

CHAPTER 6

GENERAL DISCUSSION, CONCLUSION, AND RECOMMENDATIONS

6.1 General Discussion

After reports of a continuous decline in malaria transmission for over a decade in endemic areas (WHO, 2018), there have been recent reports of increased malaria incidence and death globally, especially from *P. falciparum* (WHO 2021, 2022a). *P. falciparum* merozoites invade erythrocytes, exploring protein ligands to bind to receptors on the surface of the erythrocyte (Cowman et al., 2017). These ligand-receptor interactions during merozoite invasion set off the erythrocytic multiplication cycle that is largely responsible for the clinical signs of malaria. The parasite has a well-coordinated and sophisticated erythrocyte invasion mechanism of which the molecular details and variations across endemic settings need further investigation. Parasites are capable of switching invasion pathways probably due to environmental conditions, host specificities or interventions. Understanding how parasite strains use these different mechanisms and how these affect populations and variance in malaria incidence and outcomes could inform new control intervention approaches to combat the disease.

This study was designed to investigate variation in erythrocyte invasion pathways and the role of the molecular interactions in this process in determining malaria outcomes in The Gambia. Malaria has significantly declined in The Gambia and the main thesis posits that currently circulating parasites use the most optimal pathways to sustain infections and transmission as populations dwindle. Taking advantage of targeted deep sequencing approach, the diversity in ligands-receptor pairs in infections was investigated. The allelic distribution of glycophorin B deletions was also assessed among malaria communities with different ethnicities in The Gambia.

The first objective determined how invasion phenotypes have evolved in The Gambia over time and the effect of multiplicity of infection and ligand expression on invasion pathways. The results showed a predominant use of sialic acid (SA) independent invasion pathways, contrary to the use of sialic acid-dependent pathways previously reported over a decade ago when malaria prevalence was still high in The Gambia (Baum et al., 2003; Gomez-Escobar et al., 2010). Thus, with the decline of malaria prevalence in The Gambia, the erythrocyte invasion pathways have switched from the use of sialic acid receptors to sialic acid-independent receptors. This change from the use of SA-dependent to SA-independent pathways in currently circulating *P. falciparum* strains could be due to adaptation from the parasite to continue its life cycle in the host against pressure from intensified malaria control strategies. The thought that these changes could also be due to changes in ligand gene expression led to the selection of six ligand genes (EBA-181, EBA-175, Rh4, Rh5, Rh2b, and Clag2) for targeted transcriptomics. Varying transcript levels of the six ligand genes were observed with the Rh and EBA genes clustering separately. Interestingly, Rh genes were highly expressed in recently collected samples, probably accounting for the predominant use of SA-independent pathways of invasion. Intriguingly and contrary to expectations, the recent infections characterised were more complex, with more than a single strain in several cases. However, there was evidence of an overall decrease in heterozygosity compared to previous years. Analysis of the invasion inhibition against the dominant allele of each MSP2 allelic family (3D7 and FC27) indicated a statistically significant ($P = 0.027$) strain effect, especially for high trypsin treatment, which cleaves several receptors including GPC and GPA. This indicated that the shift to more sialic acid-independent invasion is driven by more prevalent 3D7 genomic background strains in circulation. Indeed, the overall decline in malaria prevalence in The Gambia was reflected by the decreasing heterozygosity at the MSP2 locus, despite an unexpected increase in MOI, probably driving the incidence of malaria in The Gambia in 2021 and 2022. Therefore, as

malaria declines, a smaller number of clones remain in circulation and result in low complex infections with incidence rebound. The maintenance of diversity could also be due to the importation of infection from neighbouring countries into urban centres.

The arms race between the malaria parasite and its human host also includes interactions during invasion, in which the parasite maximises invasion and growth in erythrocytes, while human populations evolve to resist the deadly erythrocytic cycle. While receptors of merozoites and merozoite ligands have mostly been studied independently, dual analysis of their variation in infections across population groups could inform how combinations are maximised and open avenues for new interventions. This could also further knowledge of the pathogenesis of the disease and a better understanding of the complex life cycle of the parasite. In chapter 4 (objective 2) of this thesis, variations in host receptor and parasite ligand genes in malaria infected individuals were investigated. To achieve this objective, 12 invasion ligands and four erythrocyte receptors were genotyped using deep amplicon sequencing on the Oxford Nanopore Technology. Single nucleotide polymorphisms (SNPs) extracted from amplicon showed ligands from different *P. falciparum* parasite infections to be mostly similar (similar genetic makeup). Similarly, variants from human receptors did not strongly group individuals by ethnicity or region of origin apart frim the Wolof ethnic group and Brikama population. Thus the parasite isolates in The Gambia have been mixing continuously, probably because of the movement of infections across the country due to population migration that promotes gene flow and recombination. The Gambia is the smallest country in mainland Africa and the short distances across sites might promote higher frequencies of parasite mixture. These findings were similar to previous reports for malaria parasites from other malaria-endemic areas in West Africa. As sampling was done across multiple sites, some stratification of human populations by variants in the receptors was expected. This was largely not observed, although variants from GPC allowed for the identification of sub-populations, driven largely by ethnicity. This

included a sub-group of individuals from the most malarious region in the country, where the Fulani population is predominant. The Fulani have mostly been reported to be protected from severe malaria, and we found SNPs in CR1 and GPC that were significantly associated with malaria phenotypes. A single variant of a parasite gene, SURFIN 4.2 shows a weak association with malaria phenotypes. Overall, the weak linking of receptors and ligands was largely missed given the poor sequence quality and the low number of variants across most isolates.

As the erythrocyte invasion process used several ligands together, it can be expected that haplotypes across the various ligands could be selected together by combinations of receptors in the population (Quintana et al., 2018; Winter et al., 2005). Thus, strong linkage disequilibrium (LD) observed between variants in some *P. falciparum* markers (DBLMSP and Surfin 4.2) and human receptors (glycophorin B, Glycophorin C, and CR1) might indicate there are several major haplotypes of ligands and receptors in circulation. This high level of haplotype diversity in ligand genes might be an adaptation mechanism to diversity in host receptors. For the human receptors, CR1 was highly polymorphic with many SNPs associated with severe malaria. Malaria has had a strong selective pressure on human populations and CR1 may carry some of these signatures. Some of the SNPs identified have not been associated with malaria previously in The Gambia and warrant further investigations.

Amongst malaria parasite receptors, the glycophorins are very important in the *Plasmodium* life cycle because they are highly used for erythrocyte invasion. The glycophorin A, B, and E (GPABE) locus on chromosome four of the human genome is highly polymorphic, containing both SNPs and structural variants like deletion, duplications, fusion, and insertions. One of these structural variants called the Dantu commonly found in East Africa has been associated with a 40 % decreased risk of developing severe malaria (Leftler et al., 2017) by increasing red blood cell tension, leading to a reduction in erythrocyte invasion rate (Kariuki et al., 2020). The most prevalent structural variant in West Africa is either complete or partial deletion of the

GPB gene. Currently, there are eight deletions on the GPB gene, numbered from the most prevalence (deletion 1) to the least prevalence (deletion 8) (Leftler et al., 2017). PCR-based genotyping assays have been developed for the two most prevalent deletions (1 and 2) enabling a cheaper and affordable means of investigating the distribution of these deletions in malaria-endemic areas in an attempt to understand their association with malaria. To this effect, a restriction fragment length polymorphism (RFLP) and multiplex PCR assays (Amuzu et al., 2021; Lane et al., 2020) were used in the third objective (chapter 5) of this study to determine the distribution of GPB DEL 1 and DEL 2 among seven ethnic groups across three regions in The Gambia, using samples collected from malaria positive cases in 2020, 2021, and 2022. The overall allele frequency of GPB DEL 1 and DEL 2 was 9.92 % and 1.7 % respectively. The Fulani ethnic group had the highest frequencies of GPB DEL 1 (14.8%). This allele frequency in the Fulani was comparable with DEL 1 (15 %) frequency observed in a previous study in West Africa using datasets from the HapMap and 1000 genome project (Leftler et al., 2017). However, the frequency of DEL 1 observed among the Jola (11.5 %) is higher than the 3.6 % previously reported in the same ethnic group. The Fulani ethnic group had previously been known to show some level of resistance against malaria (Fall et al., 2023; Troye-Blomberg et al., 2020). Thus, the high prevalence of this deletion could be part of the adaptation strategy to combat malaria in the population. On the other hand, the GPB DEL 2 variant was only present in three ethnic groups: the Serehule (2.5 %), Mandinka (2.4 %), and the Fulani (2.2 %). These findings were like those reported in the in-silico study where no DEL 2 variant was observed in the Wolof group in The Gambia. Coincidentally, the majority of individuals with these deletions were residents of the Upper River Region of the country, where malaria has the highest prevalence.

These varying DEL 1 and DEL 2 frequencies across the different ethnic groups could imply that the two variants are subject to different selective pressures. The high prevalence of DEL

1 in The Gambia might suggest a role in malaria protection, although this study was not designed to further determine if it offers protection. Moreover, DEL 1 was not associated with severe malaria, even though the analysis was poorly powered by the small number of DEL 1 variants individuals identified in the severe cases. Other genetic variants known to confer protection against malaria like the G6PD and sickle cell are selectively prevalent in malaria-endemic areas (Clarke et al., 2017; Makani et al., 2010; Uyoga et al., 2017). The results of this study suggest that the sickle cell variant and the GPB deletion variants might be linked or co-evolving in the population. This might suggest a joint effort of the variants to offer protection against malaria. Given the low sample size and the limited number of severe malaria used, further studies will be needed to investigate this hypothesis.

6.2 Conclusion

The study uses a multiprong approach (cell biology of invasion, targeted transcriptomics, and genomics) to show

1. A more complex infection in The Gambia but only of a limited number of alleles and the limited number of strains remaining in circulation are adapting to invade human erythrocytes primarily through sialic-acid independent pathways. Malaria parasites in The Gambia therefore remain resilient irrespective of intensified investigations over the last two decades.
2. That the parasite as well as the human population have been mixing continuously, with an arms race between parasite ligands and receptors resulting in significant diversity in their haplotypes in *P. falciparum* and the human population respectively. Also, some of the variants in these ligands and receptors affect the outcome of disease (susceptibility). CR1 and glycophorin C seem to have important roles in malaria outcomes that warrant further investigation.

3. The glycophorin B deletion 1 variant is highly prevalent in The Gambia human population. The frequency of glycophorin B deletion 1 was highest in the Fulani ethnic group, while the Serehule accounted for the highest frequency of glycophorin B deletion 2. Higher prevalence of these deletions in the highly malaria prevalent Upper River Region of The Gambia is suggestive of a role in malaria that should also be further investigated.

Overall, this is the first attempt to present a comprehensive molecular understanding of malaria invasion and outcomes by combining phenotyping, gene expression, and genotype of *P. falciparum* parasite and human hosts in the same population. The study have generated current knowledge, and baseline data on both the phenotypic state of the parasite and genotypic states of both the parasite and the human host in an era of very low malaria prevalence in The Gambia following intensified control interventions. These findings are important for continuous genomic surveillance of malaria as the fight towards malaria elimination in The Gambia is up-scaled.

6.3 Study Limitations

- Apart from HbS, other factors associated with malaria severity like G6PD deficiency, and alpha thalassemia were not examined.
- Only malaria-positive individuals were included in the human genotyping part of the study.

6.4 Recommendations

A. Recommendations from this Study

1. Future studies should make use of other methods of studying genetic diversity like microsattelites and sequencing approaches alongside the MSP2 used in this study. Whole transcriptomics approaches will also give a bigger picture of the transcriptional profile of *P. falciparium* during erythrocyte invasion.
2. The use of whole *P. falciparum* and human genome sequencing in the same population would be a better approach to unravel host-parasite interactions.
3. Further investigations are required to better understand the relationship between glycophorin B deletions, Hb variation, and malaria outcomes. Studies should also consider other genetic factors like alpha thalassemia, G6PD alongside all the glycophorin B deletions (deletion 1 to 8).

B. Recommendations Critical in Thesis Area

Future studies should include asymptomatic as well as malaria-negative individuals and consider a holistic analysis where all the data are analysed together to determine correlates of *P. falciparium* invasion phenotypes, host/parasite genotypes and malaria outcomes.

REFERENCES AND APENDICES

REFERENCES

Abdalla, S. H., & Pasvol, G. (2004). *Malaria: A hematological perspective*. Imperial College Press.

Adams, J. H., Blair, P. L., Kaneko, O., & Peterson, D. S. (2001). An expanding ebl family of Plasmodium falciparum. *Trends in Parasitology*, *17*(6), 297–299. https://doi.org/10.1016/S1471-4922(01)01948-1

Adams, S., Brown, H., & Turner, G. (2002). Breaking down the blood–brain barrier: Signaling a path to cerebral malaria? *Trends in Parasitology*, *18*(8), 360–366. https://doi.org/10.1016/S1471-4922(02)02353-X

Alexandre, J. S., Kaewthamasorn, M., Yahata, K., Nakazawa, S., & Kaneko, O. (2011). Positive selection on the Plasmodium falciparum clag2 gene encoding a component of the erythrocyte-binding rhoptry protein complex. *Tropical Medicine and Health*, *39*(3), 77–82. https://doi.org/10.2149/tmh.2011-12

Algady, W., Louzada, S., Carpenter, D., Brajer, P., Färnert, A., Rooth, I., Ngasala, B., Yang, F., Shaw, M.-A., & Hollox, E. J. (2018). The Malaria-Protective Human Glycophorin Structural Variant DUP4 Shows Somatic Mosaicism and Association with Hemoglobin Levels. *The American Journal of Human Genetics*, *103*(5), 769–776. https://doi.org/10.1016/j.ajhg.2018.10.008

Algady, W., Weyell, E., Mateja, D., Garcia, A., Courtin, D., & Hollox, E. J. (2021). Genotyping complex structural variation at the malaria-associated human glycophorin locus using a PCR-based strategy. *Annals of Human Genetics*, *85*(1), 7–17. https://doi.org/10.1111/ahg.12405

Amale, A., Acharya, S., Shukla, S., & Dubey, S. (2013). Falciparum Malaria Infection in a Case of Sickle Cell Trait; Unbalancing the Balanced Polymorphism. *Indian Journal of*

Hematology and Blood Transfusion, *29*(2), 123–125. https://doi.org/10.1007/s12288-012-0158-7

Amambua-Ngwa, A., Amenga-Etego, L., Kamau, E., Amato, R., Ghansah, A., Golassa, L., Randrianarivelojosia, M., Ishengoma, D., Apinjoh, T., Maïga-Ascofaré, O., Andagalu, B., Yavo, W., Bouyou-Akotet, M., Kolapo, O., Mane, K., Worwui, A., Jeffries, D., Simpson, V., D'Alessandro, U., … Djimde, A. A. (2019). Major subpopulations of *Plasmodium falciparum* in Sub-Saharan Africa. *Science*, *365*(6455), 813–816. https://doi.org/10.1126/science.aav5427

Amambua-Ngwa, A., Park, D. J., Volkman, S. K., Barnes, K. G., Bei, A. K., Lukens, A. K., Sene, P., Van Tyne, D., Ndiaye, D., Wirth, D. F., Conway, D. J., Neafsey, D. E., & Schaffner, S. F. (2012). SNP Genotyping Identifies New Signatures of Selection in a Deep Sample of West African Plasmodium falciparum Malaria Parasites. *Molecular Biology and Evolution*, *29*(11), 3249–3253. https://doi.org/10.1093/molbev/mss151

Amuzu, D. S., Rockett, K. A., Leffler, E. M., Ansah, F., Amoako, N., Morang'a, C. M., Hubbart, C., Rowlands, K., Jeffreys, A. E., Amenga-Etego, L. N., Kwiatkowski, D. P., & Awandare, G. A. (2021). High-throughput genotyping assays for identification of glycophorin B deletion variants in population studies. *Experimental Biology and Medicine*, *246*(8), 916–928. https://doi.org/10.1177/1535370220968545

Archer, N. M., Petersen, N., Clark, M. A., Buckee, C. O., Childs, L. M., & Duraisingh, M. T. (2018). Resistance to *Plasmodium falciparum* in sickle cell trait erythrocytes is driven by oxygen-dependent growth inhibition. *Proceedings of the National Academy of Sciences*, *115*(28), 7350–7355. https://doi.org/10.1073/pnas.1804388115

Ashley, E. A., & Phyo, A. P. (2018). Drugs in Development for Malaria. *Drugs*, *78*(9), 861–879. https://doi.org/10.1007/s40265-018-0911-9

Ashline, D. J., Duk, M., Lukasiewicz, J., Reinhold, V. N., Lisowska, E., & Jaskiewicz, E. (2015). The structures of glycophorin C N-glycans, a putative component of the GPC receptor site for Plasmodium falciparum EBA-140 ligand. *Glycobiology*, *25*(5), 570–581. https://doi.org/10.1093/glycob/cwu188

Atuh, N. I., Anong, D. N., Jerome, F.-C., Oriero, E., Mohammed, N. I., D'Alessandro, U., & Amambua-Ngwa, A. (2022). High genetic complexity but low relatedness in *Plasmodium falciparum* infections from Western Savannah Highlands and coastal equatorial Lowlands of Cameroon. *Pathogens and Global Health*, *116*(7), 428–437. https://doi.org/10.1080/20477724.2021.1953686

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), e2012061. https://doi.org/10.4084/mjhid.2012.061

Awandare, G. A., Nyarko, P. B., Aniweh, Y., Ayivor-Djanie, R., & Stoute, J. A. (2018). *Plasmodium falciparum* strains spontaneously switch invasion phenotype in suspension culture. *Scientific Reports*, *8*(1), 5782. https://doi.org/10.1038/s41598-018-24218-0

Baldwin, M. R., Li, X., Hanada, T., Liu, S.-C., & Chishti, A. H. (2015). Merozoite surface protein 1 recognition of host glycophorin A mediates malaria parasite invasion of red blood cells. *Blood*, *125*(17), 2704–2711. https://doi.org/10.1182/blood-2014-11-611707

Baldwin, M., Yamodo, I., Ranjan, R., Li, X., Mines, G., Marinkovic, M., Hanada, T., Oh, S. S., & Chishti, A. H. (2014). Human erythrocyte band 3 functions as a receptor for the sialic acid-independent invasion of Plasmodium falciparum. Role of the RhopH3–MSP1 complex. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1843*(12), 2855–2870. https://doi.org/10.1016/j.bbamcr.2014.08.008

Bancells, C., Llorà-Batlle, O., Poran, A., Nötzel, C., Rovira-Graells, N., Elemento, O., Kafsack, B. F. C., & Cortés, A. (2018). Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*. *Nature Microbiology*, *4*(1), 144–154. https://doi.org/10.1038/s41564-018-0291-7

Bartholdson, S. J., Bustamante, L. Y., Crosnier, C., Johnson, S., Lea, S., Rayner, J. C., & Wright, G. J. (2012). Semaphorin-7A Is an Erythrocyte Receptor for *P. falciparum* Merozoite-Specific TRAP Homolog, MTRAP. *PLoS Pathogens*, *8*(11), e1003031. https://doi.org/10.1371/journal.ppat.1003031

Battle, K. E., & Baird, J. K. (2021). The global burden of Plasmodium vivax malaria is obscure and insidious. *PLOS Medicine*, *18*(10), e1003799. https://doi.org/10.1371/journal.pmed.1003799

Baum, J., Chen, L., Healer, J., Lopaticki, S., Boyle, M., Triglia, T., Ehlgen, F., Ralph, S. A., Beeson, J. G., & Cowman, A. F. (2009). Reticulocyte-binding protein homologue 5 – An essential adhesin involved in invasion of human erythrocytes by *Plasmodium falciparum*. *International Journal for Parasitology*, *39*(3), 371–380. https://doi.org/10.1016/j.ijpara.2008.10.006

Baum, J., Maier, A. G., Good, R. T., Simpson, K. M., & Cowman, A. F. (2005). Invasion by *P. falciparum* Merozoites Suggests a Hierarchy of Molecular Interactions. *PLoS Pathogens*, *1*(4), e37. https://doi.org/10.1371/journal.ppat.0010037

Baum, J., Pinder, M., & Conway, D. J. (2003). Erythrocyte Invasion Phenotypes of *Plasmodium falciparum* in The Gambia. *Infection and Immunity*, *71*(4), 1856–1863. https://doi.org/10.1128/IAI.71.4.1856-1863.2003

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.-W., Green, J. L., Holder, A. A., & Cowman, A. F. (2006). A Conserved Molecular Motor Drives Cell Invasion and Gliding Motility across Malaria Life Cycle Stages and Other Apicomplexan

Parasites. *Journal of Biological Chemistry*, *281*(8), 5197–5208. https://doi.org/10.1074/jbc.M509807200

Baum, J., Thomas, A. W., & Conway, D. J. (2003). Evidence for Diversifying Selection on Erythrocyte-Binding Antigens of *Plasmodium falciparum* and *P. vivax*. *Genetics*, *163*(4), 1327–1336. https://doi.org/10.1093/genetics/163.4.1327

Bechtsi, D. P., & Waters, A. P. (2017). Genomics and epigenetics of sexual commitment in Plasmodium. *International Journal for Parasitology*, *47*(7), 425–434. https://doi.org/10.1016/j.ijpara.2017.03.002

Beeson, J. G., Drew, D. R., Boyle, M. J., Feng, G., Fowkes, F. J. I., & Richards, J. S. (2016). Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS Microbiology Reviews*, *40*(3), 343–372. https://doi.org/10.1093/femsre/fuw001

Bei, A. K., & Duraisingh, M. T. (2012). Functional analysis of erythrocyte determinants of Plasmodium infection. *International Journal for Parasitology*, *42*(6), 575–582. https://doi.org/10.1016/j.ijpara.2012.03.006

Bei, A. K., & Duraisingh, M. T. (2015). Measuring Plasmodium falciparum Erythrocyte Invasion Phenotypes Using Flow Cytometry. In A. Vaughan (Ed.), *Malaria Vaccines* (Vol. 1325, pp. 167–186). Springer New York. https://doi.org/10.1007/978-1-4939-2815-6_14

Benegal, V. (2005). India: Alcohol and public health. *Addiction*, *100*(8), 1051–1056. https://doi.org/10.1111/j.1360-0443.2005.01176.x

Besteiro, S., Michelin, A., Poncet, J., Dubremetz, J.-F., & Lebrun, M. (2009). Export of a Toxoplasma gondii Rhoptry Neck Protein Complex at the Host Cell Membrane to Form the Moving Junction during Invasion. *PLoS Pathogens*, *5*(2), e1000309. https://doi.org/10.1371/journal.ppat.1000309

Bigham, A. W., Magnaye, K., Dunn, D. M., Weiss, R. B., & Bamshad, M. (2018). Complex signatures of natural selection at GYPA. *Human Genetics*, *137*(2), 151–160. https://doi.org/10.1007/s00439-018-1866-3

Blair, P. L. (2002). Transcripts of developmentally regulated Plasmodium falciparum genes quantified by real-time RT-PCR. *Nucleic Acids Research*, *30*(10), 2224–2231. https://doi.org/10.1093/nar/30.10.2224

Blumenfeld, O.-O., & Huang, C.-H. (1997). Molecular genetics of glycophorin MNS variants. *Transfusion Clinique et Biologique*, *4*(4), 357–365. https://doi.org/10.1016/S1246-7820(97)80041-9

Bøgh, C., Clarke, S. E., Jawara, M., Thomas, C. J., & Lindsay, S. W. (2003). Localized breeding of the *Anopheles gambiae* complex (Diptera: Culicidae) along the River Gambia, West Africa. *Bulletin of Entomological Research*, *93*(4), 279–287. https://doi.org/10.1079/BER2003239

Boushab, B. M., Ould Ahmedou Salem, M. S., Ould Mohamed Salem Boukhary, A., Parola, P., & Basco, L. (2020). Clinical Features and Mortality Associated with Severe Malaria in Adults in Southern Mauritania. *Tropical Medicine and Infectious Disease*, *6*(1), 1. https://doi.org/10.3390/tropicalmed6010001

Bowyer, P. W., Stewart, L. B., Aspeling-Jones, H., Mensah-Brown, H. E., Ahouidi, A. D., Amambua-Ngwa, A., Awandare, G. A., & Conway, D. J. (2015). Variation in Plasmodium falciparum Erythrocyte Invasion Phenotypes and Merozoite Ligand Gene Expression across Different Populations in Areas of Malaria Endemicity. *Infection and Immunity*, *83*(6), 2575–2582. https://doi.org/10.1128/IAI.03009-14

Bozdech, Z., Llinás, M., Pulliam, B. L., Wong, E. D., Zhu, J., & DeRisi, J. L. (2003). The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum. *PLoS Biology*, *1*(1), e5. https://doi.org/10.1371/journal.pbio.0000005

Bruneel, F., Hocqueloux, L., Alberti, C., Wolff, M., Chevret, S., Bédos, J.-P., Durand, R., Le Bras, J., Régnier, B., & Vachon, F. (2003). The Clinical Spectrum of Severe Imported Falciparum Malaria in the Intensive Care Unit: Report of 188 Cases in Adults. *American Journal of Respiratory and Critical Care Medicine*, *167*(5), 684–689. https://doi.org/10.1164/rccm.200206-631OC

Buffet, P. A., Safeukui, I., Milon, G., Mercereau-Puijalon, O., & David, P. H. (2009). Retention of erythrocytes in the spleen: A double-edged process in human malaria: *Current Opinion in Hematology*, *16*(3), 157–164. https://doi.org/10.1097/MOH.0b013e32832a1d4b

Bunn, H. F. (2013). The triumph of good over evil: Protection by the sickle gene against malaria. *Blood*, *121*(1), 20–25. https://doi.org/10.1182/blood-2012-08-449397

Bunyaratvej, A., Butthep, P., Yuthavong, Y., Fucharoen, S., Khusmith, S., Yoksan, S., & Wasi, P. (1986). Increased Phagocytosis of *Plasmodium falciparum*-Infected Erythrocytes with Haemoglobin E by Peripheral Blood Monocytes. *Acta Haematologica*, *76*(2–3), 155–158. https://doi.org/10.1159/000206041

Campeotto, I., Galaway, F., Mehmood, S., Barfod, L. K., Quinkert, D., Kotraiah, V., Phares, T. W., Wright, K. E., Snijders, A. P., Draper, S. J., Higgins, M. K., & Wright, G. J. (2020). The Structure of the Cysteine-Rich Domain of Plasmodium falciparum P113 Identifies the Location of the RH5 Binding Site. *mBio*, *11*(5), e01566-20. https://doi.org/10.1128/mBio.01566-20

Camus, D., Hadley, T.J., 1985. A *Plasmodium falciparum* antigen that binds to host erythrocytes and merozoites. *Science*. 230, 553–556

Carlson, J., Helmby, H., Wahlgren, M., Carlson, J., Helmby, H., Wahlgren, M., Hill, Avs., Brewster, D., & Greenwood, B. M. (1990). Human cerebral malaria: Association with

erythrocyte rosetting and lack of anti-rosetting antibodies. *The Lancet*, *336*(8729), 1457–1460. https://doi.org/10.1016/0140-6736(90)93174-N

Caroca, C., & De Lima, J. P. (2016). Sickle Cell Trait, Malaria and Sensorineural Hearing Loss–A Case-Control Study from São Tomé and Príncipe. *Otolaryngology*, *06*(06). https://doi.org/10.4172/2161-119X.1000278

CDC, C. for D. C. and P. (2022, March 22). *CDC - Malaria—About Malaria—Disease*. https://www.cdc.gov/malaria/about/disease.html

Ceesay, S. J., Casals-Pascual, C., Erskine, J., Anya, S. E., Duah, N. O., Fulford, A. J., Sesay, S. S., Abubakar, I., Dunyo, S., Sey, O., Palmer, A., Fofana, M., Corrah, T., Bojang, K. A., Whittle, H. C., Greenwood, B. M., & Conway, D. J. (2008). Changes in malaria indices between 1999 and 2007 in The Gambia: A retrospective analysis. *The Lancet*, *372*(9649), 1545–1554. https://doi.org/10.1016/S0140-6736(08)61654-2

Ceesay, S. J., Casals-Pascual, C., Nwakanma, D. C., Walther, M., Gomez-Escobar, N., Fulford, A. J. C., Takem, E. N., Nogaro, S., Bojang, K. A., Corrah, T., Jaye, M. C., Taal, M. A., Sonko, A. A. J., & Conway, D. J. (2010). Continued Decline of Malaria in The Gambia with Implications for Elimination. *PLoS ONE*, *5*(8), e12242. https://doi.org/10.1371/journal.pone.0012242

Chasis, J., & Mohandas, N. (1992). Red blood cell glycophorins. *Blood*, *80*(8), 1869–1879. https://doi.org/10.1182/blood.V80.8.1869.1869

Chen, L., Xu, Y., Wong, W., Thompson, J. K., Healer, J., Goddard-Borger, E. D., Lawrence, M. C., & Cowman, A. F. (2017). Structural basis for inhibition of erythrocyte invasion by antibodies to Plasmodium falciparum protein CyRPA. *eLife*, *6*, e21347. https://doi.org/10.7554/eLife.21347

Chiu, C. Y. H., Healer, J., Thompson, J. K., Chen, L., Kaul, A., Savergave, L., Raghuwanshi, A., Li Wai Suen, C. S. N., Siba, P. M., Schofield, L., Mueller, I., Cowman, A. F., &

Hansen, D. S. (2014). Association of antibodies to Plasmodium falciparum reticulocyte binding protein homolog 5 with protection from clinical malaria. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00314

Chotivanich, K., Udomsangpetch, R., Pattanapanyasat, K., Chierakul, W., Simpson, J., Looareesuwan, S., & White, N. (2002). Hemoglobin E: A balanced polymorphism protective against high parasitemias and thus severe P falciparum malaria. *Blood*, *100*(4), 1172–1176.

Clarke, G. M., Rockett, K., Kivinen, K., Hubbart, C., Jeffreys, A. E., Rowlands, K., Jallow, M., Conway, D. J., Bojang, K. A., Pinder, M., Usen, S., Sisay-Joof, F., Sirugo, G., Toure, O., Thera, M. A., Konate, S., Sissoko, S., Niangaly, A., Poudiougou, B., … MalariaGEN Consortium. (2017). Characterisation of the opposing effects of G6PD deficiency on cerebral malaria and severe malarial anaemia. *eLife*, *6*, e15085. https://doi.org/10.7554/eLife.15085

Cochran, W.G. (1977) Sampling Techniques. 3rd Edition, John Wiley & Sons, New York.

Coetzee, M., & Fontenille, D. (2004). Advances in the study of Anopheles funestus, a major vector of malaria in Africa. *Insect Biochemistry and Molecular Biology*, *34*(7), 599–605. https://doi.org/10.1016/j.ibmb.2004.03.012

Coetzee, M., Hunt, R. H., Wilkerson, R., Torre, A. D., Coulibaly, M. B., & Besansky, N. J. (2013). Anopheles coluzzii and Anopheles amharicus, new members of the Anopheles gambiae complex. *Zootaxa*, *3619*(3), 246–274. https://doi.org/10.11646/zootaxa.3619.3.2

Cohuet, A., Simard, F., Toto, J.-C., Kengne, P., Coetzee, M., & Fontenille, D. (2003). Species identification within the Anopheles funestus group of malaria vectors in Cameroon and evidence for a new species. *The American Journal of Tropical Medicine and Hygiene*, *69*(2), 200–205.

Conway, D. J., Greenwood, B. M., & McBride, J. S. (1992). Longitudinal study of Plasmodium falciparum polymorphic antigens in a malaria-endemic population. *Infection and Immunity*, *60*(3), 1122–1127. https://doi.org/10.1128/iai.60.3.1122-1127.1992

Cortés, A. (2008). Switching Plasmodium falciparum genes on and off for erythrocyte invasion. *Trends in Parasitology*, *24*(11), 517–524. https://doi.org/10.1016/j.pt.2008.08.005

Cosio, F. G., Shen, X. P., Birmingham, D. J., Van Aman, M., & Hebert, L. A. (1990). Evaluation of the mechanisms responsible for the reduction in erythrocyte complement receptors when immune complexes form in vivo in primates. *Journal of Immunology (Baltimore, Md.: 1950)*, *145*(12), 4198–4206.

Cowman, A. F., & Crabb, B. S. (2006). Invasion of Red Blood Cells by Malaria Parasites. *Cell*, *124*(4), 755–766. https://doi.org/10.1016/j.cell.2006.02.006

Cowman, A. F., Berry, D., & Baum, J. (2012). The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *The Journal of Cell Biology,* 198(6), 961–971. doi:10.1083/jcb.201206112

Cowman, A. F., Healer, J., Marapana, D., & Marsh, K. (2016). Malaria: Biology and Disease. *Cell*, *167*(3), 610–624. https://doi.org/10.1016/j.cell.2016.07.055

Cowman, A. F., Tonkin, C. J., Tham, W.-H., & Duraisingh, M. T. (2017). The Molecular Basis of Erythrocyte Invasion by Malaria Parasites. *Cell Host & Microbe*, *22*(2), 232–245. https://doi.org/10.1016/j.chom.2017.07.003

Crosnier, C., Bustamante, L. Y., Bartholdson, S. J., Bei, A. K., Theron, M., Uchikawa, M., Mboup, S., Ndir, O., Kwiatkowski, D. P., Duraisingh, M. T., Rayner, J. C., & Wright, G. J. (2011). Basigin is a receptor essential for erythrocyte invasion by Plasmodium falciparum. *Nature*, *480*(7378), 534–537. https://doi.org/10.1038/nature10606

Dankwa, S., Chaand, M., Kanjee, U., Jiang, R. H. Y., Nobre, L. V., Goldberg, J. M., Bei, A. K., Moechtar, M. A., Grüring, C., Ahouidi, A. D., Ndiaye, D., Dieye, T. N., Mboup, S., Weekes, M. P., & Duraisingh, M. T. (2017). Genetic Evidence for Erythrocyte Receptor Glycophorin B Expression Levels Defining a Dominant Plasmodium falciparum Invasion Pathway into Human Erythrocytes. *Infection and Immunity*, *85*(10), e00074-17. https://doi.org/10.1128/IAI.00074-17

De Bruin, D., Lanzer, M., & Ravetch, J. V. (1994). The polymorphic subtelomeric regions of Plasmodium falciparum chromosomes contain arrays of repetitive sequence elements. *Proceedings of the National Academy of Sciences*, *91*(2), 619–623. https://doi.org/10.1073/pnas.91.2.619

De Souza, B. J., Hafalla, J. C. R., Riley, E. M., & Couper, K. N. (2010). Cerebral malaria: Why experimental murine models are required to understand the pathogenesis of disease. *Parasitology*, *137*(5), 755–772. https://doi.org/10.1017/S0031182009991715

Deans, A.-M., Nery, S., Conway, D. J., Kai, O., Marsh, K., & Rowe, J. A. (2007). Invasion Pathways and Malaria Severity in Kenyan *Plasmodium falciparum* Clinical Isolates. *Infection and Immunity*, *75*(6), 3014–3020. https://doi.org/10.1128/IAI.00249-07

DeSimone, T. M., Bei, A. K., Jennings, C. V., & Duraisingh, M. T. (2009). Genetic analysis of the cytoplasmic domain of the PfRh2b merozoite invasion protein of Plasmodium falciparum. *International Journal for Parasitology*, *39*(4), 399–405. https://doi.org/10.1016/j.ijpara.2008.08.008

Dluzewski, A. R., Ling, I. T., Hopkins, J. M., Grainger, M., Margos, G., Mitchell, G. H., Holder, A. A., & Bannister, L. H. (2008). Formation of the Food Vacuole in Plasmodium falciparum: A Potential Role for the 19 kDa Fragment of Merozoite Surface Protein 1 (MSP119). *PLoS ONE*, *3*(8), e3085. https://doi.org/10.1371/journal.pone.0003085

Dolan, S. A., Miller, L. H., & Wellems, T. E. (1990). Evidence for a switching mechanism in the invasion of erythrocytes by Plasmodium falciparum. *Journal of Clinical Investigation*, *86*(2), 618–624. https://doi.org/10.1172/JCI114753

Douglas, A. D., Baldeviano, G. C., Lucas, C. M., Lugo-Roman, L. A., Crosnier, C., Bartholdson, S. J., Diouf, A., Miura, K., Lambert, L. E., Ventocilla, J. A., Leiva, K. P., Milne, K. H., Illingworth, J. J., Spencer, A. J., Hjerrild, K. A., Alanine, D. G. W., Turner, A. V., Moorhead, J. T., Edgel, K. A., … Draper, S. J. (2015). A PfRH5-Based Vaccine Is Efficacious against Heterologous Strain Blood-Stage Plasmodium falciparum Infection in Aotus Monkeys. *Cell Host & Microbe*, *17*(1), 130–139. https://doi.org/10.1016/j.chom.2014.11.017

Douglas, A. D., Williams, A. R., Illingworth, J. J., Kamuyu, G., Biswas, S., Goodman, A. L., Wyllie, D. H., Crosnier, C., Miura, K., Wright, G. J., Long, C. A., Osier, F. H., Marsh, K., Turner, A. V., Hill, A. V. S., & Draper, S. J. (2011). The blood-stage malaria antigen PfRH5 is susceptible to vaccine-inducible cross-strain neutralizing antibody. *Nature Communications*, *2*(1), 601. https://doi.org/10.1038/ncomms1615

Doumbo, K.O., Thera, A.M., Koné, K.A., Raza, A., Tempest, J.L., Lyke, E.K., & Plowe, V.C. (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), 987–993. https://doi.org/10.4269/ajtmh.2009.09-0406

Driss, A., Hibbert, J. M., Wilson, N. O., Iqbal, S. A., Adamkiewicz, T. V., & Stiles, J. K. (2011). Genetic polymorphisms linked to susceptibility to malaria. *Malaria Journal*, *10*(1), 271. https://doi.org/10.1186/1475-2875-10-271

Duraisingh, M. T., Maier, A. G., Triglia, T., & Cowman, A. F. (2003). Erythrocyte-binding antigen 175 mediates invasion in *Plasmodium falciparum* utilizing sialic acid-

dependent and -independent pathways. *Proceedings of the National Academy of Sciences*, *100*(8), 4796–4801. https://doi.org/10.1073/pnas.0730883100

Duraisingh, M. T., Triglia, T., Ralph, S. A., Rayner, J. C., Barnwell, J. W., McFadden, G. I., & Cowman, A. F. (2003). Phenotypic variation of Plasmodium falciparum merozoite proteins directs receptor targeting for invasion of human erythrocytes. *The EMBO Journal*, *22*(5), 1047–1057. https://doi.org/10.1093/emboj/cdg096

Fairhurst, R. M., Baruch, D. I., Brittain, N. J., Ostera, G. R., Wallach, J. S., Hoang, H. L., Hayton, K., Guindo, A., Makobongo, M. O., Schwartz, O. M., Tounkara, A., Doumbo, O. K., Diallo, D. A., Fujioka, H., Ho, M., & Wellems, T. E. (2005). Abnormal display of PfEMP-1 on erythrocytes carrying haemoglobin C may protect against malaria. *Nature*, *435*(7045), 1117–1121. https://doi.org/10.1038/nature03631

Fairhurst, R. M., Fujioka, H., Hayton, K., Collins, K. F., & Wellems, T. E. (2003). Aberrant development of Plasmodium falciparum in hemoglobin CC red cells: Implications for the malaria protective effect of the homozygous state. *Blood*, *101*(8), 3309–3315. https://doi.org/10.1182/blood-2002-10-3105

Fall, A. K. D. J., Kana, I. H., Garcia-Senosiain, A., Henry, B., Dechavanne, C., Garcia, A., Buffet, P., Sabbagh, A., Migot-Nabias, F., Theisen, M., & Courtin, D. (2023). Susceptibility to malaria in fulani, Bariba, Otamari and gando individuals living in sympatry in Benin: Role of opsonizing antibodies to Plasmodium falciparum merozoites. *Heliyon*, *9*(1), e13092. https://doi.org/10.1016/j.heliyon.2023.e13092

Farashi, S., & Harteveld, C. L. (2018). Molecular basis of α-thalassemia. *Blood Cells, Molecules, and Diseases*, *70*, 43–53. https://doi.org/10.1016/j.bcmd.2017.09.004

Fenton, B., Clark, J. T., Khan, C. M., Robinson, J. V., Walliker, D., Ridley, R., et al. (1991). Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface

antigen (MSA-2) of the malaria parasite *Plasmodium falciparum*. *Molecular and Cellular Biology, 11*(2), 963–971. https://doi.org/10.1128/MCB.11.2.963

Filby, A., Begum, J., Jalal, M., & Day, W. (2015). Appraising the suitability of succinimidyl and lipophilic fluorescent dyes to track proliferation in non-quiescent cells by dye dilution. *Methods*, *82*, 29–37. https://doi.org/10.1016/j.ymeth.2015.02.016

Foko, L. P., Kumar, A., Hawadak, J., & Singh, V. (2022). Plasmodium cynomolgi in humans: Current knowledge and future directions of an emerging zoonotic malaria parasite. *Infection*, *51*(3), 623–640. https://doi.org/10.1007/s15010-022-01952-2

Fola, A. A., Abby Harrison, G. L., Hazairin, M. H., Barnadas, C., Hetzel, M. W., Iga, J., Siba, P. M., Mueller, I., & Barry, A. E. (2017). Higher Complexity of Infection and Genetic Diversity of *Plasmodium vivax* Than *Plasmodium falciparum* across all Malaria Transmission Zones of Papua New Guinea. *The American Journal of Tropical Medicine and Hygiene*, 16–0716. https://doi.org/10.4269/ajtmh.16-0716

Foth, B. J., & McFadden, G. I. (2003). The apicoplast: A plastid in Plasmodium falciparum and other apicomplexan parasites. In *International Review of Cytology* (Vol. 224, pp. 57–110). Elsevier. https://doi.org/10.1016/S0074-7696(05)24003-2

Fowkes, F. J. I., Allen, S. J., Allen, A., Alpers, M. P., Weatherall, D. J., & Day, K. P. (2008). Increased Microerythrocyte Count in Homozygous α+-Thalassaemia Contributes to Protection against Severe Malarial Anaemia. *PLoS Medicine*, *5*(3), e56. https://doi.org/10.1371/journal.pmed.0050056

Fucharoen, S., & Weatherall, D. J. (2012). The Hemoglobin E Thalassemias. *Cold Spring Harbor Perspectives in Medicine*, *2*(8), a011734–a011734. https://doi.org/10.1101/cshperspect.a011734

Fuehrer, H.-P., & Noedl, H. (2014). Recent Advances in Detection of Plasmodium ovale: Implications of Separation into the Two Species Plasmodium ovale wallikeri and

Plasmodium ovale curtisi. *Journal of Clinical Microbiology*, *52*(2), 387–391. https://doi.org/10.1128/JCM.02760-13

Galaway, F., Drought, L. G., Fala, M., Cross, N., Kemp, A. C., Rayner, J. C., & Wright, G. J. (2017). P113 is a merozoite surface protein that binds the N terminus of Plasmodium falciparum RH5. *Nature Communications*, *8*(1), 14333. https://doi.org/10.1038/ncomms14333

Gambia population. Worldmeters. https://www.worldometers.info/worldpopulation/gambiapopulation/. Access 04/08/2024.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M.-S., Nene, V., … Barrell, B. (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature*, *419*(6906), 498–511. https://doi.org/10.1038/nature01097

Gaur, D., Mayer, D. C. G., & Miller, L. H. (2004). Parasite ligand–host receptor interactions during invasion of erythrocytes by Plasmodium merozoites. *International Journal for Parasitology*, *34*(13–14), 1413–1429. https://doi.org/10.1016/j.ijpara.2004.10.010

Gaur, D., Singh, S., Singh, S., Jiang, L., Diouf, A., & Miller, L. H. (2007). Recombinant *Plasmodium falciparum* reticulocyte homology protein 4 binds to erythrocytes and blocks invasion. *Proceedings of the National Academy of Sciences*, *104*(45), 17789–17794. https://doi.org/10.1073/pnas.0708772104

Genton, B., Betuela, I., Felger, I., Al-Yaman, F., Anders, R. F., Saul, A., Rare, L., Baisor, M., Lorry, K., Brown, G. V., Pye, D., Irving, D. O., Smith, T. A., Beck, H., & Alpers, M. P. (2002). A Recombinant Blood-Stage Malaria Vaccine Reduces *Plasmodium falciparum* Density and Exerts Selective Pressure on Parasite Populations in a Phase 1–

2b Trial in Papua New Guinea. *The Journal of Infectious Diseases*, *185*(6), 820–827. https://doi.org/10.1086/339342

Gilberger, T.-W., Thompson, J. K., Triglia, T., Good, R. T., Duraisingh, M. T., & Cowman, A. F. (2003). A Novel Erythrocyte Binding Antigen-175 Paralogue fromPlasmodium falciparum Defines a New Trypsin-resistant Receptor on Human Erythrocytes. *Journal of Biological Chemistry*, *278*(16), 14480–14486. https://doi.org/10.1074/jbc.M211446200

GitHub.Human_PF_Corelation:Rscripts.GitHub. https://github.com/MPBmrcg/Human_PF_Corelation/tree/main/Rscripts

Githui, E. K., Peterson, D. S., Aman, R. A., & Abdi, A. I. (2010). Prevalence of 5′ insertion mutants and analysis of single nucleotide polymorphism in the erythrocyte binding-like 1 (ebl-1) gene in Kenyan Plasmodium falciparum field isolates. *Infection, Genetics and Evolution*, *10*(6), 833–838. https://doi.org/10.1016/j.meegid.2009.10.005

Goel, S., Palmkvist, M., Moll, K., Joannin, N., Lara, P., R Akhouri, R., Moradi, N., Öjemalm, K., Westman, M., Angeletti, D., Kjellin, H., Lehtiö, J., Blixt, O., Ideström, L., Gahmberg, C. G., Storry, J. R., Hult, A. K., Olsson, M. L., Von Heijne, G., … Wahlgren, M. (2015). RIFINs are adhesins implicated in severe Plasmodium falciparum malaria. *Nature Medicine*, *21*(4), 314–317. https://doi.org/10.1038/nm.3812

Goel, V. K., Li, X., Chen, H., Liu, S.-C., Chishti, A. H., & Oh, S. S. (2003). Band 3 is a host receptor binding merozoite surface protein 1 during the *Plasmodium falciparum* invasion of erythrocytes. *Proceedings of the National Academy of Sciences*, *100*(9), 5164–5169. https://doi.org/10.1073/pnas.0834959100

Gomez-Escobar, N., Amambua-Ngwa, A., Walther, M., Okebe, J., Ebonyi, A., & Conway, D. J. (2010). Erythrocyte Invasion and Merozoite Ligand Gene Expression in Severe and

Mild *Plasmodium falciparum* Malaria. *The Journal of Infectious Diseases*, *201*(3), 444–452. https://doi.org/10.1086/649902

Hadley, T. J., Klotz, F. W., Pasvol, G., Haynes, J. D., McGinniss, M. H., Okubo, Y., & Miller, L. H. (1987). Falciparum malaria parasites invade erythrocytes that lack glycophorin A and B (MkMk). Strain differences indicate receptor heterogeneity and two pathways for invasion. *Journal of Clinical Investigation*, *80*(4), 1190–1193. https://doi.org/10.1172/JCI113178

Hartmeyer, G. N., Stensvold, C. R., Fabricius, T., Marmolin, E. S., Hoegh, S. V., Nielsen, H. V., Kemp, M., & Vestergaard, L. S. (2019). *Plasmodium cynomolgi* as Cause of Malaria in Tourist to Southeast Asia, 2018. *Emerging Infectious Diseases*, *25*(10), 1936–1939. https://doi.org/10.3201/eid2510.190448

Hayton, K., Gaur, D., Liu, A., Takahashi, J., Henschen, B., Singh, S., Lambert, L., Furuya, T., Bouttenot, R., Doll, M., Nawaz, F., Mu, J., Jiang, L., Miller, L. H., & Wellems, T. E. (2008). Erythrocyte Binding Protein PfRH5 Polymorphisms Determine Species-Specific Pathways of Plasmodium falciparum Invasion. *Cell Host & Microbe*, *4*(1), 40–51. https://doi.org/10.1016/j.chom.2008.06.001

Healer, J., Crawford, S., Ralph, S., McFadden, G., & Cowman, A. F. (2002). Independent Translocation of Two Micronemal Proteins in Developing *Plasmodium falciparum* Merozoites. *Infection and Immunity*, *70*(10), 5751–5758. https://doi.org/10.1128/IAI.70.10.5751-5758.2002

Hedrick, P. W. (2011). Population genetics of malaria resistance in humans. *Heredity*, *107*(4), 283–304. https://doi.org/10.1038/hdy.2011.16

Hikosaka, K., Watanabe, Y., Kobayashi, F., Waki, S., Kita, K., & Tanabe, K. (2011). Highly conserved gene arrangement of the mitochondrial genomes of 23 Plasmodium species.

Parasitology International, *60*(2), 175–180. https://doi.org/10.1016/j.parint.2011.02.001

Hollox, E. J., & Louzada, S. (2023). Genetic variation of glycophorins and infectious disease. *Immunogenetics*, *75*(3), 201–206. https://doi.org/10.1007/s00251-022-01280-7

Howes, R. E., Piel, F. B., Patil, A. P., Nyangiri, O. A., Gething, P. W., Dewi, M., Hogg, M. M., Battle, K. E., Padilla, C. D., Baird, J. K., & Hay, S. I. (2012). G6PD Deficiency Prevalence and Estimates of Affected Populations in Malaria Endemic Countries: A Geostatistical Model-Based Map. *PLoS Medicine*, *9*(11), e1001339. https://doi.org/10.1371/journal.pmed.1001339

Hu, G., Cabrera, A., Kono, M., Mok, S., Chaal, B. K., Haase, S., Engelberg, K., Cheemadan, S., Spielmann, T., Preiser, P. R., Gilberger, T.-W., & Bozdech, Z. (2010). Transcriptional profiling of growth perturbations of the human malaria parasite Plasmodium falciparum. *Nature Biotechnology*, *28*(1), 91–98. https://doi.org/10.1038/nbt.1597

Hughes, A., L. (1992). Positive selection and interallelic recombination at the merozoite surface antigen-1 (MSA-1) locus of Plasmodium falciparum. *Molecular Biology and Evolution*. https://doi.org/10.1093/oxfordjournals.molbev.a040730

Hughes, Philip, N., Lucas Starnes, G., Taylor, S., & Waters, A. P. (2010). From cradle to grave: RNA biology in malaria parasites. *Wiley Interdisciplinary Reviews: RNA*, *1*(2), 287–303. https://doi.org/10.1002/wrna.30

Hviid, L., & Jensen, A. T. R. (2015). PfEMP1 – A Parasite Protein Family of Key Importance in Plasmodium falciparum Malaria Immunity and Pathogenesis. In *Advances in Parasitology* (Vol. 88, pp. 51–84). Elsevier. https://doi.org/10.1016/bs.apar.2015.02.004

Illingworth, J. J., Alanine, D. G., Brown, R., Marshall, J. M., Bartlett, H. E., Silk, S. E., Labbé, G. M., Quinkert, D., Cho, J. S., Wendler, J. P., Pattinson, D. J., Barfod, L., Douglas, A. D., Shea, M. W., Wright, K. E., de Cassan, S. C., Higgins, M. K., & Draper, S. J. (2019). Functional Comparison of Blood-Stage Plasmodium falciparum Malaria Vaccine Candidate Antigens. *Frontiers in Immunology*, *10*, 1254. https://doi.org/10.3389/fimmu.2019.01254

Ingram, V. M. (1958). Abnormal human haemoglobins. *Biochimica et Biophysica Acta*, *28*, 539–545. https://doi.org/10.1016/0006-3002(58)90516-X

Iriko, H., Kaneko, O., Otsuki, H., Tsuboi, T., Su, X., Tanabe, K., & Torii, M. (2008). Diversity and evolution of the rhoph1/clag multigene family of Plasmodium falciparum. *Molecular and Biochemical Parasitology*, *158*(1), 11–21. https://doi.org/10.1016/j.molbiopara.2007.11.004

Jallow, M., Teo, Y., Small, K. *et al.* (2009). Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nat Genet*, **41**, 657–665. https://doi.org/10.1038/ng.388

Janse, C. J., Ponnudurai, T., Lensen, A. H. W., Meuwissen, J. H. E. Th., Ramesar, J., Van Der Ploeg, M., & Overdulve, J. P. (1988). DNA synthesis in gametocytes of *Plasmodium falciparum*. *Parasitology*, *96*(1), 1–7. https://doi.org/10.1017/S0031182000081609

Jaskiewicz, E., Peyrard, T., Kaczmarek, R., Zerka, A., Jodlowska, M., & Czerwinski, M. (2018). The Gerbich blood group system: Old knowledge, new importance. *Transfusion Medicine Reviews*, *32*(2), 111–116. https://doi.org/10.1016/j.tmrv.2018.02.004

Jensen, R.T.A, Magistrado, P., Sharp, S., Joergensen, L., Lavstsen, T., Chiucchiuini, A., Salanti, A., Vestergaard, S.L., Lusingu, P.J., Hermsen, R., Sauerwein, R., Christensen, J., Nielsen, A.M., Hviid, L., Sutherland, C., Staalsoe, T., & Theander, G.T. (2004). *Plasmodium falciparum* associated with severe childhood malaria preferentially

expresses PfEMP1 encoded by group a *var* genes. *J Exp Med* (2004) 199 (9): 1179–1190. https://doi.org/10.1084/jem.20040274

Jeyaprakasam, N. K., Liew, J. W. K., Low, V. L., Wan-Sulaiman, W.-Y., & Vythilingam, I. (2020). Plasmodium knowlesi infecting humans in Southeast Asia: What's next? *PLOS Neglected Tropical Diseases*, *14*(12), e0008900. https://doi.org/10.1371/journal.pntd.0008900

Jin, J., Tarrant, R. D., Bolam, E. J., Angell-Manning, P., Soegaard, M., Pattinson, D. J., Dulal, P., Silk, S. E., Marshall, J. M., Dabbs, R. A., Nugent, F. L., Barrett, J. R., Hjerrild, K. A., Poulsen, L., Jørgensen, T., Brenner, T., Baleanu, I. N., Parracho, H. M., Tahiri-Alaoui, A., … Draper, S. J. (2018). Production, quality control, stability, and potency of cGMP-produced Plasmodium falciparum RH5.1 protein vaccine expressed in Drosophila S2 cells. *Npj Vaccines*, *3*(1), 32. https://doi.org/10.1038/s41541-018-0071-7

Johansson, E., & B. Falk, A. (2021). Erythrocytes as Biomarkers of Virus and Bacteria in View of Metal Ion Homeostasis. In K. Fatima Shad (Ed.), *Erythrocyte—A Peripheral Biomarker For Infection and Inflammation*. IntechOpen. https://doi.org/10.5772/intechopen.97850

Johnson, K. E., & Voight, B. F. (2018). Patterns of shared signatures of recent positive selection across human populations. *Nature Ecology & Evolution*, *2*(4), 713–720. https://doi.org/10.1038/s41559-018-0478-6

Josling, G. A., Petter, M., Oehring, S. C., Gupta, A. P., Dietz, O., Wilson, D. W., Schubert, T., Längst, G., Gilson, P. R., Crabb, B. S., Moes, S., Jenoe, P., Lim, S. W., Brown, G. V., Bozdech, Z., Voss, T. S., & Duffy, M. F. (2015). A Plasmodium Falciparum Bromodomain Protein Regulates Invasion Gene Expression. *Cell Host & Microbe*, *17*(6), 741–751. https://doi.org/10.1016/j.chom.2015.05.009

Kaewthamasorn, M., Yahata, K., Alexandre, J. S. F., Xangsayarath, P., Nakazawa, S., Torii, M., Sattabongkot, J., Udomsangpetch, R., & Kaneko, O. (2012). Stable allele frequency distribution of the polymorphic region of SURFIN4.2 in Plasmodium falciparum isolates from Thailand. *Parasitology International*, *61*(2), 317–323. https://doi.org/10.1016/j.parint.2011.12.003

Kafsack, B. F. C., Rovira-Graells, N., Clark, T. G., Bancells, C., Crowley, V. M., Campino, S. G., Williams, A. E., Drought, L. G., Kwiatkowski, D. P., Baker, D. A., Cortés, A., & Llinás, M. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature*, *507*(7491), 248–252. https://doi.org/10.1038/nature12920

Kaneko, O., Mu, J., Tsuboi, T., Su, X., Torii, M., 2002. Gene structure and expression of a *Plasmodium falciparum* 220-kDa protein homologous to the *Plasmodium vivax* reticulocyte binding proteins. *Mol. Biochem. Parasitol*. 121, 275–278

Kariuki, S. N., & Williams, T. N. (2020). Human genetics and malaria resistance. *Human Genetics*, *139*(6–7), 801–811. https://doi.org/10.1007/s00439-020-02142-6

Kariuki, S. N., Marin-Menendez, A., Introini, V., Ravenhill, B. J., Lin, Y.-C., Macharia, A., Makale, J., Tendwa, M., Nyamu, W., Kotar, J., Carrasquilla, M., Rowe, J. A., Rockett, K., Kwiatkowski, D., Weekes, M. P., Cicuta, P., Williams, T. N., & Rayner, J. C. (2020). Red blood cell tension protects against severe malaria in the Dantu blood group. *Nature*, *585*(7826), 579–583. https://doi.org/10.1038/s41586-020-2726-6

Kengne, P., Antonio-Nkondjio, C., Awono-Ambene, H. P., Simard, F., Awolola, T. S., & Fontenille, D. (2007). Molecular differentiation of three closely related members of the mosquito species complex, Anopheles moucheti, by mitochondrial and ribosomal DNA polymorphism. *Medical and Veterinary Entomology*, *21*(2), 177–182. https://doi.org/10.1111/j.1365-2915.2007.00681.x

Kengne, P., Awono-Ambene, P., Nkondjio, C. A., Simard, F., & Fontenille, D. (2003). Molecular identification of the Anopheles nili group of African malaria vectors. *Medical and Veterinary Entomology*, *17*(1), 67–74. https://doi.org/10.1046/j.1365-2915.2003.00411.x

Kessler, A., Dankwa, S., Bernabeu, M., Harawa, V., Danziger, S. A., Duffy, F., Kampondeni, S. D., Potchen, M. J., Dambrauskas, N., Vigdorovich, V., Oliver, B. G., Hochman, S. E., Mowrey, W. B., MacCormick, I. J. C., Mandala, W. L., Rogerson, S. J., Sather, D. N., Aitchison, J. D., Taylor, T. E., … Kim, K. (2017). Linking EPCR-Binding PfEMP1 to Brain Swelling in Pediatric Cerebral Malaria. *Cell Host & Microbe*, *22*(5), 601-614.e5. https://doi.org/10.1016/j.chom.2017.09.009

Kim, C., Colin, Y., Blanchard, D., Dahr, W., London, J., & Cartron, J.-P. (1987). Gerbich blood group deficiency of the Ge:-1,-2,-3 and Ge:-1,-2,3 types.. Immunochemical study and genomic analysis with cDNA probes. *European Journal of Biochemistry*, *165*(3), 571–579. https://doi.org/10.1111/j.1432-1033.1987.tb11478.x

Ko, W.-Y., Kaercher, K. A., Giombini, E., Marcatili, P., Froment, A., Ibrahim, M., Lema, G., Nyambo, T. B., Omar, S. A., Wambebe, C., Ranciaro, A., Hirbo, J. B., & Tishkoff, S. A. (2011). Effects of Natural Selection and Gene Conversion on the Evolution of Human Glycophorins Coding for MNS Blood Polymorphisms in Malaria-Endemic African Populations. *The American Journal of Human Genetics*, *88*(6), 741–754. https://doi.org/10.1016/j.ajhg.2011.05.005

Kobayashi, I. (1992). Mechanisms for gene conversion and homologous recombination: The double-strand break repair model and the successive half crossing-over model. *Advances in Biophysics*, *28*, 81–133. https://doi.org/10.1016/0065-227X(92)90023-K

Kohne, E. (2011). Hemoglobinopathies. *Deutsches Ärzteblatt International*. https://doi.org/10.3238/arztebl.2011.0532

Kremer, K., Kamin, D., Rittweger, E., Wilkes, J., Flammer, H., Mahler, S., Heng, J., Tonkin, C. J., Langsley, G., Hell, S. W., Carruthers, V. B., Ferguson, D. J. P., & Meissner, M. (2013). An Overexpression Screen of Toxoplasma gondii Rab-GTPases Reveals Distinct Transport Routes to the Micronemes. *PLoS Pathogens*, *9*(3), e1003213. https://doi.org/10.1371/journal.ppat.1003213

Kreuels, B., Kreuzberg, C., Kobbe, R., Ayim-Akonor, M., Apiah-Thompson, P., Thompson, B., Ehmen, C., Adjei, S., Langefeld, I., Adjei, O., & May, J. (2010). Differing effects of HbS and HbC traits on uncomplicated falciparum malaria, anemia, and child growth. *Blood*, *115*(22), 4551–4558. https://doi.org/10.1182/blood-2009-09-241844

Kussell, E., & Leibler, S. (2005). Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments. *Science*, *309*(5743), 2075–2078. https://doi.org/10.1126/science.1114383

Kwiatkowski, D. P. (2005). How malaria has affected the human genome and what human genetics can teach us about malaria. *American Journal of Human Genetics*, *77*(2), 171–192. https://doi.org/10.1086/432519

Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., Dubremetz, J.-F., Fauquenoy, S., Tomavo, S., Faber, B. W., Kocken, C. H., Thomas, A. W., Boulanger, M. J., Bentley, G. A., & Lebrun, M. (2011). The RON2-AMA1 Interaction is a Critical Step in Moving Junction-Dependent Invasion by Apicomplexan Parasites. *PLoS Pathogens*, *7*(2), e1001276. https://doi.org/10.1371/journal.ppat.1001276

Lane, W. J., Gleadall, N. S., Aeschlimann, J., Vege, S., Sanchis-Juan, A., Stephens, J., Cone Sullivan, J., Mah, H. H., Aguad, M., Smeland-Wagman, R., Lebo, M. S., Vijay Kumar, P. K., Kaufman, R. M., Green, R. C., Ouwehand, W. H., & Westhoff, C. M. (2020).

Multiple *GYPB* gene deletions associated with the U− phenotype in those of African ancestry. *Transfusion*, *60*(6), 1294–1307. https://doi.org/10.1111/trf.15839

Lau, C. K. Y., Turner, L., Jespersen, J. S., Lowe, E. D., Petersen, B., Wang, C. W., Petersen, J. E. V., Lusingu, J., Theander, T. G., Lavstsen, T., & Higgins, M. K. (2015). Structural Conservation Despite Huge Sequence Diversity Allows EPCR Binding by the PfEMP1 Family Implicated in Severe Childhood Malaria. *Cell Host & Microbe*, *17*(1), 118–129. https://doi.org/10.1016/j.chom.2014.11.007

Leclerc, M. C., Hugot, J. P., Durand, P., & Renaud, F. (2004). Evolutionary relationships between 15 *Plasmodium* species from New and Old World primates (including humans): A 18S rDNA cladistic analysis. *Parasitology*, *129*(6), 677–684. https://doi.org/10.1017/S0031182004006146

Lee, W.-C., Russell, B., & Rénia, L. (2019). Sticking for a Cause: The Falciparum Malaria Parasites Cytoadherence Paradigm. *Frontiers in Immunology*, *10*, 1444. https://doi.org/10.3389/fimmu.2019.01444

Lefèvre, T., Gouagna, L.-C., Dabiré, K. R., Elguero, E., Fontenille, D., Renaud, F., Costantini, C., & Thomas, F. (2010). Beer Consumption Increases Human Attractiveness to Malaria Mosquitoes. *PLoS ONE*, *5*(3), e9546. https://doi.org/10.1371/journal.pone.0009546

Leffler, E. M., Band, G., Busby, G. B. J., Kivinen, K., Le, Q. S., Clarke, G. M., Bojang, K. A., Conway, D. J., Jallow, M., Sisay-Joof, F., Bougouma, E. C., Mangano, V. D., Modiano, D., Sirima, S. B., Achidi, E., Apinjoh, T. O., Marsh, K., Ndila, C. M., Peshu, N., … Malaria Genomic Epidemiology Network. (2017). Resistance to malaria through structural variation of red blood cell invasion receptors. *Science*, *356*(6343), eaam6393. https://doi.org/10.1126/science.aam6393

Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. https://doi.org/10.1111/2041-210X.12410

Lelliott, P. M., McMorran, B. J., Foote, S. J., & Burgio, G. (2015). The influence of host genetics on erythrocytes and malaria infection: Is there therapeutic potential? *Malaria Journal*, *14*(1), 289. https://doi.org/10.1186/s12936-015-0809-x

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. https://doi.org/10.1016/j.molbiopara.2012.01.002

Lin, C. S., Uboldi, A. D., Marapana, D., Czabotar, P. E., Epp, C., Bujard, H., Taylor, N. L., Perugini, M. A., Hodder, A. N., & Cowman, A. F. (2014). The Merozoite Surface Protein 1 Complex Is a Platform for Binding to Human Erythrocytes by Plasmodium falciparum. *Journal of Biological Chemistry*, *289*(37), 25655–25669. https://doi.org/10.1074/jbc.M114.586495

Lisowska, E. (1988). Antigenic Properties of Human Erythrocyte Glycophorins. In A. M. Wu & L. G. Adams (Eds.), *The Molecular Immunology of Complex Carbohydrates* (Vol. 228, pp. 265–315). Springer US. https://doi.org/10.1007/978-1-4613-1663-3_10

Lobo, C.-A., De Frazao, K., Rodriguez, M., Reid, M., Zalis, M., & Lustigman, S. (2004). Invasion Profiles of Brazilian Field Isolates of *Plasmodium falciparum*: Phenotypic and Genotypic Analyses. *Infection and Immunity*, *72*(10), 5886–5891. https://doi.org/10.1128/IAI.72.10.5886-5891.2004

Lobo, C.-A., Rodriguez, M., Reid, M., & Lustigman, S. (2003). Glycophorin C is the receptor for the Plasmodium falciparum erythrocyte binding ligand PfEBP-2 (baebl). *Blood*, *101*(11), 4628–4631. https://doi.org/10.1182/blood-2002-10-3076

Longo-Pendy, N. M., Boundenga, L., Kutomy, P. O. O., Mbou-Boutambe, C., Makanga, B., Moukodoum, N., Obame-Nkoghe, J., Makouloutou, P. N., Mounioko, F., Akone-Ella, R., Nkoghe-Nkoghe, L. C., Ngangue Salamba, M. F., Lekana-Douki, J. B., & Kengne, P. (2022). Systematic Review on Diversity and Distribution of Anopheles Species in Gabon: A Fresh Look at the Potential Malaria Vectors and Perspectives. *Pathogens*, *11*(6), 668. https://doi.org/10.3390/pathogens11060668

Lopez, G. H., Hyland, C. A., & Flower, R. L. (2021). Glycophorins and the MNS blood group system: A narrative review. *Annals of Blood*, *6*, 39–39. https://doi.org/10.21037/aob-21-9

Lorenz, A., & Mpaulo, S. J. (2022). Gene conversion: A non-Mendelian process integral to meiotic recombination. *Heredity*, *129*(1), 56–63. https://doi.org/10.1038/s41437-022-00523-3

Louzada, S., Algady, W., Weyell, E., Zuccherato, L. W., Brajer, P., Almalki, F., Scliar, M. O., Naslavsky, M. S., Yamamoto, G. L., Duarte, Y. A. O., Passos-Bueno, M. R., Zatz, M., Yang, F., & Hollox, E. J. (2020). Structural variation of the malaria-associated human glycophorin A-B-E region. *BMC Genomics*, *21*(1), 446. https://doi.org/10.1186/s12864-020-06849-8

Lu, X. M., Batugedara, G., Lee, M., Prudhomme, J., Bunnik, E. M., & Le Roch, K. G. (2017). Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite Plasmodium falciparum. *Nucleic Acids Research*, *45*(13), 7825–7840. https://doi.org/10.1093/nar/gkx464

Luzzatto, L. (2012). SICKLE CELL ANAEMIA AND MALARIA. *Mediterranean Journal of Hematology and Infectious Diseases*, *4*(1), e2012065. https://doi.org/10.4084/mjhid.2012.065

Luzzatto, L., & Arese, P. (2018). Favism and Glucose-6-Phosphate Dehydrogenase Deficiency. *New England Journal of Medicine*, *378*(1), 60–71. https://doi.org/10.1056/NEJMra1708111

Lyon, J. A., Angov, E., Fay, M. P., Sullivan, J. S., Girourd, A. S., Robinson, S. J., Bergmann-Leitner, E. S., Duncan, E. H., Darko, C. A., Collins, W. E., Long, C. A., & Barnwell, J. W. (2008). Protection Induced by Plasmodium falciparum MSP142 Is Strain-Specific, Antigen and Adjuvant Dependent, and Correlates with Antibody Responses. *PLoS ONE*, *3*(7), e2830. https://doi.org/10.1371/journal.pone.0002830

Mackinnon, M. J., & Marsh, K. (2010). The Selection Landscape of Malaria Parasites. *Science*, *328*(5980), 866–871. https://doi.org/10.1126/science.1185410

Maier, A. G., Baum, J., Smith, B., Conway, D. J., & Cowman, A. F. (2009). Polymorphisms in Erythrocyte Binding Antigens 140 and 181 Affect Function and Binding but Not Receptor Specificity in *Plasmodium falciparum*. *Infection and Immunity*, *77*(4), 1689–1699. https://doi.org/10.1128/IAI.01331-08

Maier, A. G., Duraisingh, M. T., Reeder, J. C., Patel, S. S., Kazura, J. W., Zimmerman, P. A., & Cowman, A. F. (2003). Plasmodium falciparum erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations. *Nature Medicine*, *9*(1), 87–92. https://doi.org/10.1038/nm807

Makani, J., Komba, A. N., Cox, S. E., Oruo, J., Mwamtemi, K., Kitundu, J., Magesa, P., Rwezaula, S., Meda, E., Mgaya, J., Pallangyo, K., Okiro, E., Muturi, D., Newton, C. R., Fegan, G., Marsh, K., & Williams, T. N. (2010). Malaria in patients with sickle cell anemia: Burden, risk factors, and outcome at the outpatient clinic and during

hospitalization. *Blood*, *115*(2), 215–220. https://doi.org/10.1182/blood-2009-07-233528

MalariaGEN. (2015). A novel locus of resistance to severe malaria in a region of ancient balancing selection. *Nature*, *526*(7572), 253–257. https://doi.org/10.1038/nature15390

Malpede, B. M., Lin, D. H., & Tolia, N. H. (2013). Molecular Basis for Sialic Acid-dependent Receptor Recognition by the Plasmodium falciparum Invasion Protein Erythrocyte-binding Antigen-140/BAEBL. *Journal of Biological Chemistry*, *288*(17), 12406–12415. https://doi.org/10.1074/jbc.M113.450643

Massey, N. C., Garrod, G., Wiebe, A., Henry, A. J., Huang, Z., Moyes, C. L., & Sinka, M. E. (2016). A global bionomic database for the dominant vectors of human malaria. *Scientific Data*, *3*(1), 160014. https://doi.org/10.1038/sdata.2016.14

Mattei, M. G., Colin, Y., Le Van Kim, C., Mattei, J. F., & Cartron, J. P. (1986). Localization of the gene for human erythrocyte glycophorin C to chromosome 2, q14–q21. *Human Genetics*, *74*(4), 420–422. https://doi.org/10.1007/BF00280497

Mayengue, P. I., Ndounga, M., Davy, M. M., Tandou, N., & Ntoumi, F. (2005). In vivo chloroquine resistance and prevalence of the pfcrt codon 76 mutation in Plasmodium falciparum isolates from the Republic of Congo. *Acta Tropica*, *95*(3), 219–225. https://doi.org/10.1016/j.actatropica.2005.06.001

Mayer, D. C. G., Cofie, J., Jiang, L., Hartl, D. L., Tracy, E., Kabat, J., Mendoza, L. H., & Miller, L. H. (2009). Glycophorin B is the erythrocyte receptor of *Plasmodium falciparum* erythrocyte-binding ligand, EBL-1. *Proceedings of the National Academy of Sciences*, *106*(13), 5348–5352. https://doi.org/10.1073/pnas.0900878106

Mayer, D. C. G., Mu, J.-B., Feng, X., Su, X., & Miller, L. H. (2002). Polymorphism in a *Plasmodium falciparum* Erythrocyte-binding Ligand Changes Its Receptor Specificity.

The Journal of Experimental Medicine, *196*(11), 1523–1528. https://doi.org/10.1084/jem.20020750

Mayer, D. C. G., Mu, J.-B., Kaneko, O., Duan, J., Su, X., & Miller, L. H. (2004). Polymorphism in the *Plasmodium falciparum* erythrocyte-binding ligand JESEBL/EBA-181 alters its receptor specificity. *Proceedings of the National Academy of Sciences*, *101*(8), 2518–2523. https://doi.org/10.1073/pnas.0307318101

Mbanefo, E. C., Ahmed, A. M., Titouna, A., Elmaraezy, A., Trang, N. T. H., Phuoc Long, N., Hoang Anh, N., Diem Nghi, T., The Hung, B., Van Hieu, M., Ky Anh, N., Huy, N. T., & Hirayama, K. (2017). Association of glucose-6-phosphate dehydrogenase deficiency and malaria: A systematic review and meta-analysis. *Scientific Reports*, *7*(1), 45963. https://doi.org/10.1038/srep45963

McCullough, J. (2014). RBCs as targets of infection. *Hematology* American Society of Hematology *Education Program*, (1): 404–409. https://doi.org/10.1182/asheducation-2014.1.404

McQuaid, F., & Rowe, J. A. (2019). Rosetting revisited: a critical look at the evidence for host erythrocyte receptors in Plasmodium falciparum rosetting. *Parasitology,* 1–43. doi:10.1017/s0031182019001288

Ménard, R., Sultan, A. A., Cortes, C., Altszuler, R., van Dijk, M. R., Janse, C. J., Waters, A. P., Nussenzweig, R. S., & Nussenzweig, V. (1997). Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes. *Nature,* **385**, 336–340 (1997). https://doi.org/10.1038/385336a0

Mensah-Brown, H. E., Amoako, N., Abugri, J., Stewart, L. B., Agongo, G., Dickson, E. K., Ofori, M. F., Stoute, J. A., Conway, D. J., & Awandare, G. A. (2015). Analysis of Erythrocyte Invasion Mechanisms of *Plasmodium falciparum* Clinical Isolates Across

3 Malaria-Endemic Areas in Ghana. *Journal of Infectious Diseases*, *212*(8), 1288–1297. https://doi.org/10.1093/infdis/jiv207

Miller, L. H., Mason, S. J., Clyde, D. F., & McGinniss, M. H. (1976). The Resistance Factor to *Plasmodium vivax* in Blacks: The Duffy-Blood-Group Genotype, *FyFy*. *New England Journal of Medicine*, *295*(6), 302–304. https://doi.org/10.1056/NEJM197608052950602

Milner, D. A. (2010). Rethinking cerebral malaria pathology. *Current Opinion in Infectious Diseases, 23(5), 456–463.* doi:10.1097/qco.0b013e32833c3dbe

Mkumbaye, S. I., Wang, C. W., Lyimo, E., Jespersen, J. S., Manjurano, A., Mosha, J., Kavishe, R. A., Mwakalinga, S. B., Minja, D. T. R., Lusingu, J. P., Theander, T. G., & Lavstsen, T. (2017). The Severity of Plasmodium falciparum Infection Is Associated with Transcript Levels of *var* Genes Encoding Endothelial Protein C Receptor-Binding *P. falciparum* Erythrocyte Membrane Protein 1. *Infection and Immunity*, *85*(4), e00841-16. https://doi.org/10.1128/IAI.00841-16

Mockenhaupt, F. P., Ehrhardt, S., Gellert, S., Otchwemah, R. N., Dietz, E., Anemana, S. D., & Bienzle, U. (2004). A+-thalassemia protects African children from severe malaria. *Blood*, *104*(7), 2003–2006. https://doi.org/10.1182/blood-2003-11-4090

Modiano, D., Luoni, G., Sirima, B. S., Simporé, J., Verra, F., Konaté, A., Rastrelli, E., Olivieri, A., Calissano, C., Paganotti, G. M., D'Urbano, L., Sanou, I., Sawadogo, A., Modiano, G., & Coluzzi, M. (2001). Haemoglobin C protects against clinical Plasmodium falciparum malaria. *Nature*, *414*(6861), 305–308. https://doi.org/10.1038/35104556

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. doi: 10.3389/fcimb.2022.816574

Morahan, B. J., Wang, L., & Coppel, R. L. (2009). No TRAP, no invasion. *Trends in Parasitology*, *25*(2), 77–84. https://doi.org/10.1016/j.pt.2008.11.004

Mota, M. M., Pradel, G., Vanderberg, J. P., Hafalla, J. C. R., Frevert, U., Nussenzweig, R. S., Nussenzweig, V., & Rodríguez, A. (2001). Migration of *Plasmodium* Sporozoites Through Cells Before Infection. *Science*, *291*(5501), 141–144. https://doi.org/10.1126/science.291.5501.141

Mwaiswelo, R. O., Mawala, W., Iversen, P. O., De Montalembert, M., Luzzatto, L., & Makani, J. (2020). Sickle cell disease and malaria: Decreased exposure and asplenia can modulate the risk from Plasmodium falciparum. *Malaria Journal*, *19*(1), 165. https://doi.org/10.1186/s12936-020-03212-w

Mwesigwa, J., Achan, J., Di Tanna, G. L., Affara, M., Jawara, M., Worwui, A., Hamid-Adiamoh, M., Kanuteh, F., Ceesay, S., Bousema, T., Drakeley, C., Grietens, K. P., Lindsay, S. W., Van Geertruyden, J.-P., & D'Alessandro, U. (2017). Residual malaria transmission dynamics varies across The Gambia despite high coverage of control interventions. *PLOS ONE*, *12*(11), e0187059. https://doi.org/10.1371/journal.pone.0187059

Narum, D. L., Fuhrmann, S. R., Luu, T., & Sim, B. K. L. (2002). A novel Plasmodium falciparum erythrocyte binding protein-2 (EBP2/BAEBL) involved in erythrocyte receptor binding. *Molecular and Biochemical Parasitology*, *119*(2), 159–168. https://doi.org/10.1016/S0166-6851(01)00428-5

Ndiath, M. O. (2019). Insecticides and Insecticide Resistance. In F. Ariey, F. Gay, & R. Ménard (Eds.), *Malaria Control and Elimination* (Vol. 2013, pp. 287–304). Springer New York. https://doi.org/10.1007/978-1-4939-9550-9_18

Nei, M. (1978). ESTIMATION OF AVERAGE HETEROZYGOSITY AND GENETIC DISTANCE FROM A SMALL NUMBER OF INDIVIDUALS. *Genetics*, *89*(3), 583–590. https://doi.org/10.1093/genetics/89.3.583

Nery, S., Deans, A.-M., Mosobo, M., Marsh, K., Rowe, J. A., & Conway, D. J. (2006). Expression of Plasmodium falciparum genes involved in erythrocyte invasion varies among isolates cultured directly from patients. *Molecular and Biochemical Parasitology*, *149*(2), 208–215. https://doi.org/10.1016/j.molbiopara.2006.05.014

Nguitragool, W., Bokhari, A. A. B., Pillai, A. D., Rayavara, K., Sharma, P., Turpin, B., Aravind, L., & Desai, S. A. (2011). Malaria Parasite clag3 Genes Determine Channel-Mediated Nutrient Uptake by Infected Red Blood Cells. *Cell*, *145*(5), 665–677. https://doi.org/10.1016/j.cell.2011.05.002

Niang, M., Bei, A. K., Madnani, K. G., Pelly, S., Dankwa, S., Kanjee, U., Gunalan, K., Amaladoss, A., Yeo, K. P., Bob, N. S., Malleret, B., Duraisingh, M. T., & Preiser, P. R. (2014). STEVOR Is a Plasmodium falciparum Erythrocyte Binding Protein that Mediates Merozoite Invasion and Rosetting. *Cell Host & Microbe*, *16*(1), 81–93. https://doi.org/10.1016/j.chom.2014.06.004

Nixon, C. P. (2016). Plasmodium falciparum gametocyte transit through the cutaneous microvasculature: A new target for malaria transmission blocking vaccines? *Human Vaccines & Immunotherapeutics*, *12*(12), 3189–3195. https://doi.org/10.1080/21645515.2016.1183076

O'Donnell, A., Premawardhena, A., Arambepola, M., Samaranayake, R., Allen, S. J., Peto, T. E. A., Fisher, C. A., Cook, J., Corran, P. H., Olivieri, N. F., & Weatherall, D. J. (2009). Interaction of malaria with a common form of severe thalassemia in an Asian population. *Proceedings of the National Academy of Sciences*, *106*(44), 18716–18721. https://doi.org/10.1073/pnas.0910142106

Ochola-Oyier, L. I., Okombo, J., Wagatua, N., Ochieng, J., Tetteh, K. K., Fegan, G., Bejon, P., & Marsh, K. (2016). Comparison of allele frequencies of Plasmodium falciparum merozoite antigens in malaria infections sampled in different years in a Kenyan population. *Malaria Journal*, *15*(1), 261. https://doi.org/10.1186/s12936-016-1304-8

Ochola-Oyier, L. I., Wamae, K., Omedo, I., Ogola, C., Matharu, A., Musabyimana, J. P., Njogu, F. K., & Marsh, K. (2019). Few Plasmodium falciparum merozoite ligand and erythrocyte receptor pairs show evidence of balancing selection. *Infection, Genetics and Evolution*, *69*, 235–245. https://doi.org/10.1016/j.meegid.2019.02.004

Ochola, L. B., Siddondo, B. R., Ocholla, H., Nkya, S., Kimani, E. N., Williams, T. N., Makale, J. O., Liljander, A., Urban, B. C., Bull, P. C., Szestak, T., Marsh, K., & Craig, A. G. (2011). Specific Receptor Usage in Plasmodium falciparum Cytoadherence Is Associated with Disease Outcome. *PLoS ONE*, *6*(3), e14741. https://doi.org/10.1371/journal.pone.0014741

Ochola, L. I., Tetteh, K. K. A., Stewart, L. B., Riitho, V., Marsh, K., & Conway, D. J. (2010). Allele Frequency-Based and Polymorphism-Versus-Divergence Indices of Balancing Selection in a New Filtered Set of Polymorphic Genes in Plasmodium falciparum. *Molecular Biology and Evolution*, *27*(10), 2344–2351. https://doi.org/10.1093/molbev/msq119

Okoyeh, J. N., Pillai, C. R., & Chitnis, C. E. (1999). *Plasmodium falciparum* Field Isolates Commonly Use Erythrocyte Invasion Pathways That Are Independent of Sialic Acid Residues of Glycophorin A. *Infection and Immunity*, *67*(11), 5784–5791. https://doi.org/10.1128/IAI.67.11.5784-5791.1999

Onda, M., Kudo, S., & Fukuda, M. (1994). Genomic organization of glycophorin A gene family revealed by yeast artificial chromosomes containing human genomic DNA. *The Journal of Biological Chemistry*, *269*(17), 13013–13020.

Onda, M., Kudo, S., Rearden, A., Mattei, M. G., & Fukuda, M. (1993). Identification of a precursor genomic segment that provided a sequence unique to glycophorin B and E genes. *Proceedings of the National Academy of Sciences*, *90*(15), 7220–7224. https://doi.org/10.1073/pnas.90.15.7220

Ord, R. L., Rodriguez, M., & Lobo, C. A. (2015). Malaria invasion ligand RH5 and its prime candidacy in blood-stage malaria vaccine design. *Human Vaccines & Immunotherapeutics*, *11*(6), 1465–1473. https://doi.org/10.1080/21645515.2015.1026496

Ord, R. L., Rodriguez, M., Yamasaki, T., Takeo, S., Tsuboi, T., & Lobo, C. A. (2012). Targeting Sialic Acid Dependent and Independent Pathways of Invasion in Plasmodium falciparum. *PLoS ONE*, *7*(1), e30251. https://doi.org/10.1371/journal.pone.0030251

Otieno, R. O., Ouma, C., Ong'echa, J. M., Keller, C. C., Were, T., Waindi, E. N., Michaels, M. G., Day, R. D., Vulule, J. M., & Perkins, D. J. (2006). Increased severe anemia in HIV-1-exposed and HIV-1-positive infants and children during acute malaria. *AIDS*, *20*(2), 275–280. https://doi.org/10.1097/01.aids.0000200533.56490.b7

Ouattara, A., Takala-Harrison, S., Thera, M. A., Coulibaly, D., Niangaly, A., Saye, R., Tolo, Y., Dutta, S., Heppner, D. G., Soisson, L., Diggs, C. L., Vekemans, J., Cohen, J., Blackwelder, W. C., Dube, T., Laurens, M. B., Doumbo, O. K., & Plowe, C. V. (2013). Molecular Basis of Allele-Specific Efficacy of a Blood-Stage Malaria Vaccine: Vaccine Development Implications. *The Journal of Infectious Diseases*, *207*(3), 511–519. https://doi.org/10.1093/infdis/jis709

Paing, M. M., Salinas, N. D., Adams, Y., Oksman, A., Jensen, A. T., Goldberg, D. E., & Tolia, N. H. (2018). Shed EBA-175 mediates red blood cell clustering that enhances malaria parasite growth and enables immune evasion. *eLife*, *7*, e43224. https://doi.org/10.7554/eLife.43224

Palacajornsuk, P. (2006). Review: Molecular basis of MNS blood group variants. *Immunohematology*, *22*(4), 171–182.

Perkins, M. E., & Holt, E. H. (1988). Erythrocyte receptor recognition varies in Plasmodium falciparum isolates. *Molecular and Biochemical Parasitology*, *27*(1), 23–34. https://doi.org/10.1016/0166-6851(88)90021-7

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. https://doi.org/10.1016/S0166-6851(99)00173-5

Peterson, D. S., Miller, L. H., & Wellems, T. E. (1995). Isolation of multiple sequences from the Plasmodium falciparum genome that encode conserved domains homologous to those in erythrocyte-binding proteins. *Proceedings of the National Academy of Sciences*, *92*(15), 7100–7104. https://doi.org/10.1073/pnas.92.15.7100

Phillips, R. E., Looareesuwan, S., Warrell, D. A., Lee, S. H., Karbwang, J., Warrell, J., White, N. J., Swasdichai, C., & Weatherall, D. J. (1986). The Importance of Anaemia in Cerebral and Uncomplicated Falciparum Malaria: Role of Complications, Dyserythropoiesis and Iron Sequestration. *QJM: An International Journal of Medicine*. https://doi.org/10.1093/oxfordjournals.qjmed.a067960

Piel, F. B., Hay, S. I., Gupta, S., Weatherall, D. J., & Williams, T. N. (2013). Global Burden of Sickle Cell Anaemia in Children under Five, 2010–2050: Modelling Based on Demographics, Excess Mortality, and Interventions. *PLoS Medicine*, *10*(7), e1001484. https://doi.org/10.1371/journal.pmed.1001484

Polin, H., Danzer, M., Reiter, A., Brisner, M., Gaszner, W., Weinberger, J., & Gabriel, C. (2014). MN typing discrepancies based on *GYPA-B-A* hybrid. *Vox Sanguinis*, *107*(4), 393–398. https://doi.org/10.1111/vox.12168

Polley, S. D., & Conway, D. J. (2001). Strong Diversifying Selection on Domains of the *Plasmodium falciparum* Apical Membrane Antigen 1 Gene. *Genetics*, *158*(4), 1505–1512. https://doi.org/10.1093/genetics/158.4.1505

Polley, S. D., Tetteh, K. K. A., Lloyd, J. M., Akpogheneta, O. J., Greenwood, B. M., Bojang, K. A., & Conway, D. J. (2007). *Plasmodium falciparum* Merozoite Surface Protein 3 Is a Target of Allele-Specific Immunity and Alleles Are Maintained by Natural Selection. *The Journal of Infectious Diseases*, *195*(2), 279–287. https://doi.org/10.1086/509806

Poran, A., Nötzel, C., Aly, O., Mencia-Trinchant, N., Harris, C. T., Guzman, M. L., Hassane, D. C., Elemento, O., & Kafsack, B. F. C. (2017). Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature*, *551*(7678), 95–99. https://doi.org/10.1038/nature24280

Postels, G. D., Chimalizeni, F.Y., Mallewa, M., Boivin, J.M. & Seydel, B. K. (2012). Pediatric Cerebral Malaria: A Scourge of Africa. *Future Neurology*, 8(1), 67-85. https://doi.org/10.2217/fnl.12.84

Preiser, P., Kaviratne, M., Khan, S., Bannister, L., & Jarra, W. (2000). The apical organelles of malaria merozoites: Host cell selection, invasion, host immunity and immune evasion. *Microbes and Infection*, *2*(12), 1461–1477. https://doi.org/10.1016/S1286-4579(00)01301-0

Price, R. N., Simpson, J. A., Nosten, F., Luxemburger, C., Hkirjaroen, L., Kuile, F. T., Chongsuphajaisiddhi, T., & White, N. J. (2001). Factors contributing to anemia after uncomplicated falciparum malaria. *The American Journal of Tropical Medicine and Hygiene*, *65*(5), 614–622. https://doi.org/10.4269/ajtmh.2001.65.614

Quin, J. E., Bujila, I., Chérif, M., Sanou, G. S., Qu, Y., Vafa Homann, M., Rolicka, A., Sirima, S. B., O'Connell, M. A., Lennartsson, A., Troye-Blomberg, M., Nebie, I., & Östlund Farrants, A.-K. (2017). Major transcriptional changes observed in the Fulani, an ethnic group less susceptible to malaria. *eLife*, *6*, e29156. https://doi.org/10.7554/eLife.29156

Quintana, M. D. P., Ch'ng, J.-H., Zandian, A., Imam, M., Hultenby, K., Theisen, M., Nilsson, P., Qundos, U., Moll, K., Chan, S., & Wahlgren, M. (2018). SURGE complex of Plasmodium falciparum in the rhoptry-neck (SURFIN4.2-RON4-GLURP) contributes to merozoite invasion. *PLOS ONE*, *13*(8), e0201669. https://doi.org/10.1371/journal.pone.0201669

Rayner, J. C., Tran, T. M., Corredor, V., Huber, C. S., Barnwell, J. W., & Galinski, M. R. (2005). Dramatic difference in diversity between Plasmodium falciparum and Plasmodium vivax reticulocyte binding-like genes. *The American Journal of Tropical Medicine and Hygiene*, *72*(6), 666–674.

Rayner, J. C., Vargas-Serrato, E., Huber, C. S., Galinski, M. R., & Barnwell, J. W. (2001). A *Plasmodium falciparum* Homologue of *Plasmodium vivax* Reticulocyte Binding Protein (PvRBP1) Defines a Trypsin-resistant Erythrocyte Invasion Pathway. *The Journal of Experimental Medicine*, *194*(11), 1571–1582. https://doi.org/10.1084/jem.194.11.1571

Reddy, K. S., Amlabu, E., Pandey, A. K., Mitra, P., Chauhan, V. S., & Gaur, D. (2015). Multiprotein complex between the GPI-anchored CyRPA with PfRH5 and PfRipr is crucial for *Plasmodium falciparum* erythrocyte invasion. *Proceedings of the National Academy of Sciences*, *112*(4), 1179–1184. https://doi.org/10.1073/pnas.1415466112

Rees, D. C., Williams, T. N., & Gladwin, M. T. (2010). Sickle-cell disease. *The Lancet*, *376*(9757), 2018–2031. https://doi.org/10.1016/S0140-6736(10)61029-X

Regions of The Gambia. Mappr. https://www.mappr.co/countries/gambia. Access 04/08/2024.

Reid, M. E. (2009). MNS blood group system: A review. *Immunohematology*, *25*(3), 95–101.

Reiling, L., Richards, J. S., Fowkes, F. J. I., Barry, A. E., Triglia, T., Chokejindachai, W., Michon, P., Tavul, L., Siba, P. M., Cowman, A. F., Mueller, I., & Beeson, J. G. (2010). Evidence That the Erythrocyte Invasion Ligand PfRh2 is a Target of Protective Immunity against *Plasmodium falciparum* Malaria. *The Journal of Immunology*, *185*(10), 6157–6167. https://doi.org/10.4049/jimmunol.1001555

Resende, M., Nielsen, M. A., Dahlbäck, M., Ditlev, S. B., Andersen, P., Sander, A. F., Ndam, N. T., Theander, T. G., & Salanti, A. (2008). Identification of glycosaminoglycan binding regions in the Plasmodium falciparum encoded placental sequestration ligand, VAR2CSA. *Malaria Journal*, *7*(1), 104. https://doi.org/10.1186/1475-2875-7-104

Riglar, D. T., Richard, D., Wilson, D. W., Boyle, M. J., Dekiwadia, C., Turnbull, L., Angrisano, F., Marapana, D. S., Rogers, K. L., Whitchurch, C. B., Beeson, J. G., Cowman, A. F., Ralph, S. A., & Baum, J. (2011). Super-Resolution Dissection of Coordinated Events during Malaria Parasite Invasion of the Human Erythrocyte. *Cell Host & Microbe*, *9*(1), 9–20. https://doi.org/10.1016/j.chom.2010.12.003

Roh, M. E., Tessema, S. K., Murphy, M., Nhlabathi, N., Mkhonta, N., Vilakati, S., Ntshalintshali, N., Saini, M., Maphalala, G., Chen, A., Wilheim, J., Prach, L., Gosling, R., Kunene, S., S. Hsiang, M., & Greenhouse, B. (2019). High Genetic Diversity of Plasmodium falciparum in the Low-Transmission Setting of the Kingdom of Eswatini. *The Journal of Infectious Diseases*, *220*(8), 1346–1354. https://doi.org/10.1093/infdis/jiz305

Rovira-Graells, N., Gupta, A. P., Planet, E., Crowley, V. M., Mok, S., Ribas De Pouplana, L., Preiser, P. R., Bozdech, Z., & Cortés, A. (2012). Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Research*, *22*(5), 925–938. https://doi.org/10.1101/gr.129692.111

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. https://doi.org/10.1017/S1462399409001082

Rowe, J. A., Moulds, J. M., Newbold, C. I., & Miller, L. H. (1997). *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature*, *388*(6639), 292–295. https://doi.org/10.1038/40888

Rowe, J.A., Obeiro, J., Newbold, I.C., & Marsh, K. (1995). *Plasmodium falciparum* rosetting is associated with malaria severity in Kenya. *Infection and Immunity*, 63(6). https://doi.org/10.1128/iai.63.6.2323-2326.1995

Rydzak, J., Kaczmarek, R., Czerwinski, M., Lukasiewicz, J., Tyborowska, J., Szewczyk, B., & Jaskiewicz, E. (2015). The Baculovirus-Expressed Binding Region of Plasmodium falciparum EBA-140 Ligand and Its Glycophorin C Binding Specificity. *PLOS ONE*, *10*(1), e0115437. https://doi.org/10.1371/journal.pone.0115437

Sanders, P. R., Gilson, P. R., Cantin, G. T., Greenbaum, D. C., Nebl, T., Carucci, D. J., McConville, M. J., Schofield, L., Hodder, A. N., Yates, J. R., & Crabb, B. S. (2005). Distinct Protein Classes Including Novel Merozoite Surface Antigens in Raft-like Membranes of Plasmodium falciparum. *Journal of Biological Chemistry*, *280*(48), 40169–40176. https://doi.org/10.1074/jbc.M509631200

Satchwell, T. J. (2016). Erythrocyte invasion receptors for *Plasmodium falciparum*: New and old: Erythrocyte invasion receptors for *Plasmodium falciparum*. *Transfusion Medicine*, *26*(2), 77–88. https://doi.org/10.1111/tme.12280

Scherf, A., Lopez-Rubio, J. J., & Riviere, L. (2008). Antigenic Variation in *Plasmodium falciparum*. *Annual Review of Microbiology*, *62*(1), 445–470. https://doi.org/10.1146/annurev.micro.61.080706.093134

Schwartz, E., Sadetzki, S., Murad, H., & Raveh, D. (2001). Age as a Risk Factor for Severe *Plasmodium falciparum* Malaria in Nonimmune Patients. *Clinical Infectious Diseases*, *33*(10), 1774–1777. https://doi.org/10.1086/322522

Silvestrini, F., Alano, P., & Williams, J. L. (2000). Commitment to the production of male and female gametocytes in the human malaria parasite *Plasmodium falciparum*. *Parasitology*, *121*(5), 465–471. https://doi.org/10.1017/S0031182099006691

Sim, B.K., 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, 1941–1944

Sinden, R., Talman, A., Marques, S., Wass, M., & Sternberg, M. (2010). The flagellum in malarial parasites. *Current Opinion in Microbiology*, *13*(4), 491–500. https://doi.org/10.1016/j.mib.2010.05.016

Singh, S., & Chitnis, C. E. (2017). Molecular Signaling Involved in Entry and Exit of Malaria Parasites from Host Erythrocytes. *Cold Spring Harbor Perspectives in Medicine*, *7*(10), a026815. https://doi.org/10.1101/cshperspect.a026815

Singh, S., Alam, M. M., Pal-Bhowmick, I., Brzostowski, J. A., & Chitnis, C. E. (2010). Distinct External Signals Trigger Sequential Release of Apical Organelles during Erythrocyte Invasion by Malaria Parasites. *PLoS Pathogens*, *6*(2), e1000746. https://doi.org/10.1371/journal.ppat.1000746

Sisquella, X., Nebl, T., Thompson, J. K., Whitehead, L., Malpede, B. M., Salinas, N. D., Rogers, K., Tolia, N. H., Fleig, A., O'Neill, J., Tham, W.-H., David Horgen, F., & Cowman, A. F. (2017). Plasmodium falciparum ligand binding to erythrocytes induce

alterations in deformability essential for invasion. *eLife*, *6*, e21083. https://doi.org/10.7554/eLife.21083

Smith, J. D., Craig, A. G., Kriek, N., Hudson-Taylor, D., Kyes, S., Fagen, T., Pinches, R., Baruch, D. I., Newbold, C. I., & Miller, L. H. (2000). Identification of a *Plasmodium falciparum* intercellular adhesion molecule-1 binding domain: A parasite adhesion trait implicated in cerebral malaria. *Proceedings of the National Academy of Sciences*, *97*(4), 1766–1771. https://doi.org/10.1073/pnas.040545897

Snounou, G., Zhu, X., Siripoon, N., Jarra, W., Thaithong, S., Brown, K. N., & Viriyakosol, S. (1999). Biased distribution of MSP1 and MSP2 allelic variants in *Plasmodium falciparum* populations in Thailand. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *93*(4), 369–374. https://doi.org/10.1016/S0035-9203(99)90120-7

Srinivasan, P., Beatty, W. L., Diouf, A., Herrera, R., Ambroggio, X., Moch, J. K., Tyler, J. S., Narum, D. L., Pierce, S. K., Boothroyd, J. C., Haynes, J. D., & Miller, L. H. (2011). Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proceedings of the National Academy of Sciences*, *108*(32), 13275–13280. https://doi.org/10.1073/pnas.1110303108

Storm, J., Jespersen, J. S., Seydel, K. B., Szestak, T., Mbewe, M., Chisala, N. V., Phula, P., Wang, C. W., Taylor, T. E., Moxon, C. A., Lavstsen, T., & Craig, A. G. (2019). Cerebral malaria is associated with differential cytoadherence to brain endothelial cells. *EMBO Molecular Medicine*, *11*(2), e9164. https://doi.org/10.15252/emmm.201809164

Stubbs, J., Simpson, K. M., Triglia, T., Plouffe, D., Tonkin, C. J., Duraisingh, M. T., Maier, A. G., Winzeler, E. A., & Cowman, A. F. (2005). Molecular Mechanism for Switching of *P. falciparum* Invasion Pathways into Human Erythrocytes. *Science*, *309*(5739), 1384–1387. https://doi.org/10.1126/science.1115257

Sturm, A., Amino, R., Van De Sand, C., Regen, T., Retzlaff, S., Rennenberg, A., Krueger, A., Pollok, J.-M., Menard, R., & Heussler, V. T. (2006). Manipulation of Host Hepatocytes by the Malaria Parasite for Delivery into Liver Sinusoids. *Science*, *313*(5791), 1287–1290. https://doi.org/10.1126/science.1129720

Sutherland, C. J., Tanomsing, N., Nolder, D., Oguike, M., Jennison, C., Pukrittayakamee, S., Dolecek, C., Hien, T. T., Do Rosário, V. E., Arez, A. P., Pinto, J., Michon, P., Escalante, A. A., Nosten, F., Burke, M., Lee, R., Blaze, M., Otto, T. D., Barnwell, J. W., … Polley, S. D. (2010). Two Nonrecombining Sympatric Forms of the Human Malaria Parasite *Plasmodium ovale* Occur Globally. *The Journal of Infectious Diseases*, *201*(10), 1544–1550. https://doi.org/10.1086/652240

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595. https://doi.org/10.1093/genetics/123.3.585

Takala, S. L., Escalante, A. A., Branch, O. H., Kariuki, S., Biswas, S., Chaiyaroj, S. C., & Lal, A. A. (2006). Genetic diversity in the Block 2 region of the merozoite surface protein 1 (MSP-1) of Plasmodium falciparum: Additional complexity and selection and convergence in fragment size polymorphism. *Infection, Genetics and Evolution*, *6*(5), 417–424. https://doi.org/10.1016/j.meegid.2006.01.009

Tarun, A. S., Baer, K., Dumpit, R. F., Gray, S., Lejarcegui, N., Frevert, U., & Kappe, S. H. I. (2006). Quantitative isolation and in vivo imaging of malaria parasite liver stages. *International Journal for Parasitology*, *36*(12), 1283–1293. https://doi.org/10.1016/j.ijpara.2006.06.009

Taylor, H. M., Grainger, M., & Holder, A. A. (2002). Variation in the Expression of a *Plasmodium falciparum* Protein Family Implicated in Erythrocyte Invasion. *Infection and Immunity*, *70*(10), 5779–5789. https://doi.org/10.1128/IAI.70.10.5779-5789.2002

Taylor, S. M., Parobek, C. M., & Fairhurst, R. M. (2012). Haemoglobinopathies and the clinical epidemiology of malaria: A systematic review and meta-analysis. *The Lancet Infectious Diseases*, *12*(6), 457–468. https://doi.org/10.1016/S1473-3099(12)70055-5

Tham, W.-H., Healer, J., & Cowman, A. F. (2012). Erythrocyte and reticulocyte binding-like proteins of Plasmodium falciparum. *Trends in Parasitology*, *28*(1), 23–30. https://doi.org/10.1016/j.pt.2011.10.002

Tham, W.-H., Schmidt, C. Q., Hauhart, R. E., Guariento, M., Tetteh-Quarcoo, P. B., Lopaticki, S., Atkinson, J. P., Barlow, P. N., & Cowman, A. F. (2011). Plasmodium falciparum uses a key functional site in complement receptor type-1 for invasion of human erythrocytes. *Blood*, *118*(7), 1923–1933. https://doi.org/10.1182/blood-2011-03-341305

Tham, W.-H., Wilson, D. W., Lopaticki, S., Schmidt, C. Q., Tetteh-Quarcoo, P. B., Barlow, P. N., Richard, D., Corbin, J. E., Beeson, J. G., & Cowman, A. F. (2010). Complement receptor 1 is the host erythrocyte receptor for *Plasmodium falciparum* PfRh4 invasion ligand. *Proceedings of the National Academy of Sciences*, *107*(40), 17327–17332. https://doi.org/10.1073/pnas.1008151107

Tham, W.-H., Wilson, D. W., Reiling, L., Chen, L., Beeson, J. G., & Cowman, A. F. (2009). Antibodies to Reticulocyte Binding Protein-Like Homologue 4 Inhibit Invasion of *Plasmodium falciparum* into Human Erythrocytes. *Infection and Immunity*, *77*(6), 2427–2435. https://doi.org/10.1128/IAI.00048-09

Tham, W.H., Wilson, A.W., Lopaticki, S., Schmidt, C.Q., Tetteh-Quarcoo, P.B., Barlow, P.N., Richard, D., Corbin, J.E., Beeson, J. G. & Cowman, A.F. (2010). Complement receptor 1 is the host erythrocyte receptor for *Plasmodium falciparum* PfRh4 invasion ligand. 107(40): *PNAS*.17327–17332. doi: 10.1073/pnas.1008151107

Thompson JK, Triglia T, Reed MB, Cowman AF. A novel ligand from *Plasmodium*

falciparum that binds to a sialic acid-containing receptor on the surface of human erythrocytes. *Mol Biol*. 2001;41:47–58

Thathy, V., Moulds, J. M., Guyah, B., Otieno, W., & Stoute, J. A. (2005). Complement receptor 1 polymorphisms associated with resistance to severe malaria in Kenya. *Malaria Journal*, *4*(1), 54. https://doi.org/10.1186/1475-2875-4-54

Theron, M., Cross, N., Cawkill, P., Bustamante, L. Y., & Rayner, J. C. (2018). An in vitro erythrocyte preference assay reveals that Plasmodium falciparum parasites prefer Type O over Type A erythrocytes. *Scientific Reports*, *8*(1), 8133. https://doi.org/10.1038/s41598-018-26559-2

Timmann, C., Thye, T., Vens, M., Evans, J., May, J., Ehmen, C., Sievertsen, J., Muntau, B., Ruge, G., Loag, W., Ansong, D., Antwi, S., Asafo-Adjei, E., Nguah, S. B., Kwakye, K. O., Akoto, A. O. Y., Sylverken, J., Brendel, M., Schuldt, K., … Horstmann, R. D. (2012). Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature*, *489*(7416), 443–446. https://doi.org/10.1038/nature11334

Touray, A. O., Mobegi, V. A., Wamunyokoli, F., & Herren, J. K. (2020). Diversity and Multiplicity of *P. falciparum* infections among asymptomatic school children in Mbita, Western Kenya. *Scientific Reports*, *10*(1), 5924. https://doi.org/10.1038/s41598-020-62819-w

Tran, T. M., Ongoiba, A., Coursen, J., Crosnier, C., Diouf, A., Huang, C.-Y., Li, S., Doumbo, S., Doumtabe, D., Kone, Y., Bathily, A., Dia, S., Niangaly, M., Dara, C., Sangala, J., Miller, L. H., Doumbo, O. K., Kayentao, K., Long, C. A., … Crompton, P. D. (2014). Naturally Acquired Antibodies Specific for Plasmodium falciparum Reticulocyte-Binding Protein Homologue 5 Inhibit Parasite Growth and Predict Protection From Malaria. *The Journal of Infectious Diseases*, *209*(5), 789–798. https://doi.org/10.1093/infdis/jit553

Treeck, M., Zacherl, S., Herrmann, S., Cabrera, A., Kono, M., Struck, N. S., Engelberg, K., Haase, S., Frischknecht, F., Miura, K., Spielmann, T., & Gilberger, T. W. (2009). Functional Analysis of the Leading Malaria Vaccine Candidate AMA-1 Reveals an Essential Role for the Cytoplasmic Domain in the Invasion Process. *PLoS Pathogens*, *5*(3), e1000322. https://doi.org/10.1371/journal.ppat.1000322

Trenholme, K. R., Gardiner, D. L., Holt, D. C., Thomas, E. A., Cowman, A. F., & Kemp, D. J. (2000). *clag9*: A cytoadherence gene in *Plasmodium falciparum* essential for binding of parasitized erythrocytes to CD36. *Proceedings of the National Academy of Sciences*, *97*(8), 4029–4033. https://doi.org/10.1073/pnas.040561197

Treutiger, C. J., Hedlund, I., Helmby, H., Carlson, J., Jepson, A., Twuamsi, P., Kwiatkowski, D., Greenwood, B. M., & Wahlgren, M. (1992). Rosette Formation in Plasmodium falciparum Isolates and Anti-Rosette Activity of Sera from Gambians with Cerebral or Uncomplicated Malaria. *The American Journal of Tropical Medicine and Hygiene*, *46*(5), 503–510. https://doi.org/10.4269/ajtmh.1992.46.503

Triglia, T., Chen, L., Lopaticki, S., Dekiwadia, C., Riglar, D. T., Hodder, A. N., Ralph, S. A., Baum, J., & Cowman, A. F. (2011). Plasmodium falciparum Merozoite Invasion Is Inhibited by Antibodies that Target the PfRh2a and b Binding Domains. *PLoS Pathogens*, *7*(6), e1002075. https://doi.org/10.1371/journal.ppat.1002075

Triglia, T., Duraisingh, M. T., Good, R. T., & Cowman, A. F. (2004). Reticulocyte-binding protein homologue 1 is required for sialic acid-dependent invasion into human erythrocytes by Plasmodium falciparum: Function of Plasmodium erythrocyte binding proteins. *Molecular Microbiology*, *55*(1), 162–174. https://doi.org/10.1111/j.1365-2958.2004.04388.x

Triglia, T., Thompson, J., Caruana, S. R., Delorenzi, M., Speed, T., & Cowman, A. F. (2001). Identification of Proteins from *Plasmodium falciparum* That Are Homologous to

Reticulocyte Binding Proteins in *Plasmodium vivax*. *Infection and Immunity*, *69*(2), 1084–1092. https://doi.org/10.1128/IAI.69.2.1084-1092.2001

Troye-Blomberg, M., Arama, C., Quin, J., Bujila, I., & Östlund Farrants, A.-K. (2020). What will studies of Fulani individuals naturally exposed to malaria teach us about protective immunity to malaria? *Scandinavian Journal of Immunology*, *92*(4), e12932. https://doi.org/10.1111/sji.12932

Tsuneyama, H., Isa, K., Watanabe-Okochi, N., Ogasawara, K., Uchikawa, M., & Satake, M. (2020). An unusual variant glycophorin expressing protease-resistant M antigen encoded by the *GYPB-E(2-4)-B* hybrid gene. *Vox Sanguinis*, *115*(7), 579–585. https://doi.org/10.1111/vox.12918

Turner, L., Lavstsen, T., Berger, S. S., Wang, C. W., Petersen, J. E. V., Avril, M., Brazier, A. J., Freeth, J., Jespersen, J. S., Nielsen, M. A., Magistrado, P., Lusingu, J., Smith, J. D., Higgins, M. K., & Theander, T. G. (2013). Severe malaria is associated with parasite binding to endothelial protein C receptor. *Nature*, *498*(7455), 502–505. https://doi.org/10.1038/nature12216

Uyoga, S., Ndila, C. M., Macharia, A. W., Nyutu, G., Shah, S., Peshu, N., Clarke, G. M., Kwiatkowski, D. P., Rockett, K. A., & Williams, T. N. (2015). Glucose-6-phosphate dehydrogenase deficiency and the risk of malaria and other diseases in children in Kenya: A case-control and a cohort study. *The Lancet Haematology*, *2*(10), e437–e444. https://doi.org/10.1016/S2352-3026(15)00152-0

Van Tyne, D., Park, D. J., Schaffner, S. F., Neafsey, D. E., Angelino, E., Cortese, J. F., Barnes, K. G., Rosen, D. M., Lukens, A. K., Daniels, R. F., Milner, D. A., Johnson, C. A., Shlyakhter, I., Grossman, S. R., Becker, J. S., Yamins, D., Karlsson, E. K., Ndiaye, D., Sarr, O., … Sabeti, P. C. (2011). Identification and Functional Validation of the Novel

Antimalarial Resistance Locus PF10_0355 in Plasmodium falciparum. *PLoS Genetics*, *7*(4), e1001383. https://doi.org/10.1371/journal.pgen.1001383

Verra, F., Chokejindachai, W., Weedall, G. D., Polley, S. D., Mwangi, T. W., Marsh, K., & Conway, D. J. (2006). Contrasting signatures of selection on the Plasmodium falciparum erythrocyte binding antigen gene family. *Molecular and Biochemical Parasitology*, *149*(2), 182–190. https://doi.org/10.1016/j.molbiopara.2006.05.010

Verra, F., Simpore, J., Warimwe, G. M., Tetteh, K. K., Howard, T., Osier, F. H. A., Bancone, G., Avellino, P., Blot, I., Fegan, G., Bull, P. C., Williams, T. N., Conway, D. J., Marsh, K., & Modiano, D. (2007). Haemoglobin C and S Role in Acquired Immunity against Plasmodium falciparum Malaria. *PLoS ONE*, *2*(10), e978. https://doi.org/10.1371/journal.pone.0000978

Vignal, A., Rahuel, C., London, J., Cherif Zahar, B., Schaff, S., Hattab, C., Okubo, Y., & Cartron, J.-P. (1990). A novel gene member of the human glycophorin A and B gene family. Molecular cloning and expression. *European Journal of Biochemistry*, *191*(3), 619–625. https://doi.org/10.1111/j.1432-1033.1990.tb19166.x

Vogt, A. M., Pettersson, F., Moll, K., Jonsson, C., Normark, J., Ribacke, U., Egwang, T. G., Ekre, H.-P., Spillmann, D., Chen, Q., & Wahlgren, M. (2006). Release of Sequestered Malaria Parasites upon Injection of a Glycosaminoglycan. *PLoS Pathogens*, *2*(9), e100. https://doi.org/10.1371/journal.ppat.0020100

Volkman, S. K., Sabeti, P. C., DeCaprio, D., Neafsey, D. E., Schaffner, S. F., Milner, D. A., Daily, J. P., Sarr, O., Ndiaye, D., Ndir, O., Mboup, S., Duraisingh, M. T., Lukens, A., Derr, A., Stange-Thomann, N., Waggoner, S., Onofrio, R., Ziaugra, L., Mauceli, E., … Wirth, D. F. (2007). A genome-wide map of diversity in Plasmodium falciparum. *Nature Genetics*, *39*(1), 113–119. https://doi.org/10.1038/ng1930

Wanaguru, M., Liu, W., Hahn, B. H., Rayner, J. C., & Wright, G. J. (2013). RH5–Basigin interaction plays a major role in the host tropism of *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences*, *110*(51), 20735–20740. https://doi.org/10.1073/pnas.1320771110

Weatherall, D. J., & Clegg, J. B. (Eds.). (2001). *The Thalassaemia Syndromes*. Blackwell Science Ltd. https://doi.org/10.1002/9780470696705

Weiss, G. E., Crabb, B. S., & Gilson, P. R. (2016). Overlaying Molecular and Temporal Aspects of Malaria Parasite Invasion. *Trends in Parasitology*, *32*(4), 284–295. https://doi.org/10.1016/j.pt.2015.12.007

Weiss, G. E., Gilson, P. R., Taechalertpaisarn, T., Tham, W.-H., De Jong, N. W. M., Harvey, K. L., Fowkes, F. J. I., Barlow, P. N., Rayner, J. C., Wright, G. J., Cowman, A. F., & Crabb, B. S. (2015). Revealing the Sequence and Resulting Cellular Morphology of Receptor-Ligand Interactions during Plasmodium falciparum Invasion of Erythrocytes. *PLOS Pathogens*, *11*(2), e1004670. https://doi.org/10.1371/journal.ppat.1004670

WHO Regional Office for Africa. (2017, July 21). *Progress towards Malaria Elimination in The Gambia*. WHO | Regional Office for Africa. https://www.afro.who.int/news/progress-towards-malaria-elimination-gambia

WHO. (2018). *World malaria report 2019*. World Health Organization. https://www.who.int/publications/i/item/9789241565653

WHO. (2019). *World malaria report 2019*. World Health Organization. https://apps.who.int/iris/handle/10665/330011

WHO. (2020). *World malaria report 2020: 20 years of global progress and challenges*. World Health Organization. https://apps.who.int/iris/handle/10665/337660

WHO. (2021). *World malaria report 2021*. World Health Organization. https://apps.who.int/iris/handle/10665/350147

WHO. (20223). *World malaria report 2023*. World Health Organization. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2023

WHO. (2022a). *Guidelines for malaria*. World Health Organization. http://www.ncbi.nlm.nih.gov/books/NBK579022/

WHO. (2022b). *World malaria report 2022*. World Health Organization. https://www.who.int/publications/i/item/9789240064898

Wiebe, A., Longbottom, J., Gleave, K., Shearer, F. M., Sinka, M. E., Massey, N. C., Cameron, E., Bhatt, S., Gething, P. W., Hemingway, J., Smith, D. L., Coleman, M., & Moyes, C. L. (2017). Geographical distributions of African malaria vector sibling species and evidence for insecticide resistance. *Malaria Journal*, *16*(1), 85. https://doi.org/10.1186/s12936-017-1734-y

Wilder, J. A., Hewett, E. K., & Gansner, M. E. (2009). Molecular Evolution of GYPC: Evidence for Recent Structural Innovation and Positive Selection in Humans. *Molecular Biology and Evolution*, *26*(12), 2679–2687. https://doi.org/10.1093/molbev/msp183

Willemetz, A., Nataf, J., Thonier, V., Peyrard, T., & Arnaud, L. (2015). Gene conversion events between *GYPB* and *GYPE* abolish expression of the S and s blood group antigens. *Vox Sanguinis*, *108*(4), 410–416. https://doi.org/10.1111/vox.12244

Williams, T. N., Maitland, K., Bennett, S., Ganczakowski, M., Peto, T. E. A., Newbold, C. I., Bowden, D. K., Weatherall, D. J., & Clegg, J. B. (1996). High incidence of malaria in α-thalassaemic children. *Nature*, *383*(6600), 522–525. https://doi.org/10.1038/383522a0

Williams, T. N., Mwangi, T. W., Roberts, D. J., Alexander, N. D., Weatherall, D. J., Wambua, S., Kortok, M., Snow, R. W., & Marsh, K. (2005). An Immune Basis for Malaria

Protection by the Sickle Cell Trait. *PLoS Medicine*, *2*(5), e128. https://doi.org/10.1371/journal.pmed.0020128

Williams, T. N., Weatherall, D. J., & Newbold, C. I. (2002). The membrane characteristics of *Plasmodium falciparum* -infected and -uninfected heterozygous α^{0} thalassaemic erythrocytes: Malaria Protection by α Thalassaemia. *British Journal of Haematology*, *118*(2), 663–670. https://doi.org/10.1046/j.1365-2141.2002.03610.x

Winter, G., Kawai, S., Haeggström, M., Kaneko, O., Von Euler, A., Kawazu, S., Palm, D., Fernandez, V., & Wahlgren, M. (2005). SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes. *The Journal of Experimental Medicine*, *201*(11), 1853–1863. https://doi.org/10.1084/jem.20041392

Wong, W., Huang, R., Menant, S., Hong, C., Sandow, J. J., Birkinshaw, R. W., Healer, J., Hodder, A. N., Kanjee, U., Tonkin, C. J., Heckmann, D., Soroka, V., Søgaard, T. M. M., Jørgensen, T., Duraisingh, M. T., Czabotar, P. E., De Jongh, W. A., Tham, W.-H., Webb, A. I., … Cowman, A. F. (2019). Structure of Plasmodium falciparum Rh5–CyRPA–Ripr invasion complex. *Nature*, *565*(7737), 118–121. https://doi.org/10.1038/s41586-018-0779-6

Yam, X. Y., Niang, M., Madnani, K. G., & Preiser, P. R. (2017). Three Is a Crowd – New Insights into Rosetting in Plasmodium falciparum. *Trends in Parasitology*, *33*(4), 309–320. https://doi.org/10.1016/j.pt.2016.12.012

Yap, A., Azevedo, M. F., Gilson, P. R., Weiss, G. E., O'Neill, M. T., Wilson, D. W., Crabb, B. S., & Cowman, A. F. (2014). Conditional expression of apical membrane antigen 1 in *P lasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cellular Microbiology*, *16*(5), 642–656. https://doi.org/10.1111/cmi.12287

Yeoh, S., O'Donnell, R. A., Koussis, K., Dluzewski, A. R., Ansell, K. H., Osborne, S. A., Hackett, F., Withers-Martinez, C., Mitchell, G. H., Bannister, L. H., Bryans, J. S.,

Kettleborough, C. A., & Blackman, M. J. (2007). Subcellular Discharge of a Serine Protease Mediates Release of Invasive Malaria Parasites from Host Erythrocytes. *Cell*, *131*(6), 1072–1083. https://doi.org/10.1016/j.cell.2007.10.049

Yusuph, R., Sawe, H. R., Nkondora, P. N., & Mfinanga, J. A. (2019). Profile and outcomes of patients with acute complications of malaria presenting to an urban emergency department of a tertiary hospital in Tanzania. *BMC Research Notes*, *12*(1), 345. https://doi.org/10.1186/s13104-019-4388-8

Zhu, L., Van Der Pluijm, R. W., Kucharski, M., Nayak, S., Tripathi, J., White, N. J., Day, N. P. J., Faiz, A., Phyo, A. P., Amaratunga, C., Lek, D., Ashley, E. A., Nosten, F., Smithuis, F., Ginsburg, H., Von Seidlein, L., Lin, K., Imwong, M., Chotivanich, K., … Bozdech, Z. (2022). Artemisinin resistance in the malaria parasite, Plasmodium falciparum, originates from its initial transcriptional response. *Communications Biology*, *5*(1), 274. https://doi.org/10.1038/s42003-022-03215-0

Zhu, X., He, Y., Liang, Y., Kaneko, O., Cui, L., & Cao, Y. (2017). Tryptophan-rich domains of Plasmodium falciparum SURFIN4.2 and Plasmodium vivax PvSTP2 interact with membrane skeleton of red blood cell. *Malaria Journal*, *16*(1), 121. https://doi.org/10.1186/s12936-017-1772-5

Zimmerman, P.A., Fitness, J., Moulds, J.M., McNamara, D.T., Kasehagen, L.J., Alexandra Rowe, J., & Hill, A.V.S. (2003). CR1 Knops blood group alleles are not associated with severe malaria in the Gambia. *Genes & Immunity*, 4(5): 368–373. doi:10.1038/sj.gene.6363980.

APPENDICES

APPENDIX I: ETHICAL CLEARANCE

C/o MRC Unit: The Gambia @ LSHTM, Fajara
P.O. Box 273, Banjul
The Gambia, West Africa
Fax: +220 – 4495919 or 4496513
Tel: +220 – 4495442-6 Ext. 2308
Email: ethics@mrc.gm

The Gambia Government/MRC Joint
ETHICS COMMITTEE

Miss Nora Nganyewo

MRCG at LSHTM

26 November 2021

Dear Miss Nganyewo

Study Title: Genetic determinants of P. falciparum erythrocyte invasion and malaria outcome in divergent African human populations

Project ID/ethics ref: 26296

Thank you for submitting your application which was considered by the Gambia Government/MRCG Joint Ethics Committee at its meeting held on 18 November 2021.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form and supporting documentation,.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document Type	File Name	Date	Version
Investigator CV	BioSketch_NIH_AAN	01/02/2021	1
Other	Nora's Research_Ethics_online_training_certificate	16/07/2021	1
Other	Nora's Human_Tissue_online_training_Certificate_new	16/07/2021	1
Investigator CV	CV_Nganyewo Nora Nghochuzie	18/07/2021	1
Protocol / Proposal	PAMGEN-Questionnaire	03/09/2021	1
Information Sheet	PAMGEN_Information and consent FORM	03/09/2021	1
Protocol / Proposal	Nora's PhD project proposal	17/09/2021	2
Covering Letter	Ethics Clarifications Cover Letter	17/09/2021	1

After ethical review

The Principal Investigator (PI) or delegate is responsible for informing the Ethics Committee of any subsequent changes to the application. These must be submitted to the Committee for review using an Amendment form. Amendments must not be initiated before receipt of written favourable opinion from the Committee.

The PI or delegate is also required to notify the Ethics Committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reactions (SUSARs) which occur during the project by submitting a Serious Adverse Event form. An annual report should be submitted to the Committee using an Annual Report form on the anniversary of the approval of the study during the lifetime of the study. At the end of the study, the PI or delegate must notify the Committee using an End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at: http://leo.lshtm.ac.uk. Additional information is available at: www.lshtm.ac.uk/ethics.

With best wishes

Yours sincerely

Dr Mohammadou Kabir Cham

Chairperson, Gambia Government/MRCG Joint Ethics Committee

C/O MRC Unit The Gambia at LSHTM
PO Box 273 Banjul, The Gambia
West Africa
Switchboard (+220) 4495442/6 Ext 2308
Fax (+220) 4495919/4496513
E-mail: ethics@mrc.gm
Intranet: http://mrcportal/Committees/SCC/SitePages/Home.aspx
Webpage:https://mrcportal.mrc.gm/Committees/SCC/SitePages/Home.aspx

The Gambia Government/MRC Joint
ETHICS COMMITTEE

C/o MRC Unit: The Gambia, Fajara
P.O. Box 273, Banjul
The Gambia, West Africa
Fax: +220 – 4495919 or 4496513
Tel: +220 – 4495442-6 Ext. 2308
Email: ethics@mrc.gm

11 March 2014

Dr Alfred Amambua Ngwa
Disease Control and Elimination Theme
MRC Unit The Gambia
Fajara

Dear Dr Ngwa

SCC 1368v2, Plasmodium falciparum anti-malarial drug resistance in the Gambia

Thank you for submitting your proposal dated 7 February 2014 for consideration by the Gambia Government/MRC Joint Ethics Committee at its meeting held on 28 February 2014.

We are pleased to approve your proposed study.

With best wishes

Yours sincerely

Mr Malamin Sonko
Chairman, Gambia Government/MRC Joint Ethics Committee

Documents submitted for review:-

- Informed Consent Document_6 February 2014
- Informed Consent Document (malaria cohort study)_9 January 2013
- correspondence re L2008.29
- CV-Alfred Amambua Ngwa

APPENDIX II: INFORMATION AND CONSENT SHEETS

MRC Unit The Gambia

PARTICIPANT INFORMATION SHEET

Version 2.0 Date:03/09/2021

Study Title: Genetic interactions between human populations and malaria parasites in different environmental settings across Africa: Erythrocyte receptor polymorphisms and susceptibility to malaria

SCC:	1626	Protocol:	

Sponsor & Funder: Wellcome Trust / African Academy of Science

What is informed consent?

You are invited to let you or your child take part in a research study. Before you decide, you need to understand why the research study is being done and what will happen in it. Please take time to read the following information or get the information explained to you in your language. Listen carefully. You can ask questions if there is anything that is not clear or you do not understand. You may also wish to speak to your spouse, family members, friends or others before deciding to let your child take part in the study.

If you decide to allow your child to join the study, you will need to sign or put a thumbprint on a consent form saying you agree for you or your child to be in the study. You will receive a copy of the consent form.

Why is this study being done?

Parasites that infect people and cause malaria use factors in the human blood to gain entry and react to cause disease. In some people the infection can cause very complicated illness and even death, especially in young children. We now know the types of factors in the parasite and human blood that in combination will cause people to be ill or in some cases to remain well even if they have the infection. We want to study the different types of these factors that protect people from malaria in our populations in The Gambia and across different places in Africa that have different types of malaria parasite, people and the mosquitoes that transmit it. We will also look at factors in the mosquitoes that may prevent chemicals from killing mosquitoes or that have made mosquitoes to bite people outside the house.

All the information we will gather will help the government agency (National Malaria Control Programme - NMCP) to decide better ways of ensuring that malaria is eliminated from communities. We will share this information with those working in other places where peaple also suffering from malaria, so that they can carry out a similar study in future and together help in eliminating malaria from all over Africa.

What does this study involve?

If you accept or allow your child to participate, we will collect 3-5 (~75-150µl) drops of blood on filter paper from a finger prick using automated finger prick devised that will not pain. If you or your child is ill of malaria, we will take 2mL (half a teaspoon) of blood to send to the MRCG lab for more testing, including tests that can identify factors that allow malaria parasites to enter blood cells.

It is up to you to decide whether or not you or your child would take part. If you do decide to take part/ allow your child take part, you will be asked to sign or thumbprint a consent form.

What will happen to the samples taken in this study?

The samples will be taken to the MRCG laboratory at Fajara site where the following will be done;

- The first test will check for infection and how much of the malaria parasite is there. We will use other laboratory tests that can also tell us what kind of the malaria parasite is present.
- Another analysis will find the type of changes in the nature of some parasites that make them enter the blood of specific people. For this kind of tests, we will send a small quantity of the blood sample to our collaborators outside the country.
- We will apply specific genetic tests to detect changes in factors on the surface of your blood cells that affect the way parasites enter the body and cause malaria.
- We would like to keep some of the samples for future work as we continue to understand more about the parasite and how it causes malaria. So, we would also like your permission to store some of you or your child's sample at MRCG or in advanced storage freezers in other places in Africa.

We are doing the same studies in 6 other countries and all their blood samples will be sent to us at MRCG labs for similar analyses and for storage.

What harm or discomfort can you expect in the study?

You/your child will experience a slight pain from the needle/ pin when we collect the blood; this usually lasts only for a few minutes.

What benefits can you expect in the study?

There are no direct benefits to you or your child from participating in this study but the community will gain from the knowledge we will generate.

Will you be compensated for your child's/ward's participation in the study?

You and your child will not get paid by the study.

What happens if you refuse to participate in the study or change your mind later?

You or your child are free not to join the study. If we find out that you/your child are ill of malaria, you will still get the normal medical care as provided by the local health services. If we find new information during the study that may change if you/your child can be in the study, we will tell you as soon as possible.

How will your child's information be kept and who will be allowed to see it?

All information that is collected about you/your child in the study will be kept strictly confidential at MRCG under the supervision of the leader of the study. Personal information will only be seen by the study team members, the sponsor and if necessary the Ethics Committee and Government authorities. Anybody else wanting this information will seek permission from the study board at MRCG.

Who should you contact if you have questions?

If you have any questions or are worried you can call Drs Alfred Ngwa [7817428] or Joseph Okebe [7721747]. You can also call the personal numbers of the MRCG workers given to you after this, if needed.

Please feel free to ask any question you might have about the study.

Who has reviewed this study?

This study has been checked by scientists at MRCG-LSHTM and by the Gambia Government/MRCG-LSHTM Joint Ethics Committee. The Ethics Committee protects your rights and wellbeing, and has given permission for it to take place.

CONSENT / ASSENT FORM

Participant's Name _____

Participant's Identification Number: |__|__|__|__|__|__|__|__|__|__|__|__|

_____ **OR** _____
(Printed name of parent) (Printed name of guardian)

☐ I have read the written information **OR**

☐ I have had the information explained to me by study personnel in a language that I understand,

and I

- confirm that the choice for me or my child participate in this study is entirely voluntarily,
- confirm that I have had the opportunity to ask questions about this study and I am happy with the answers that have been provided,
- understand that I allow access to the information about my child by the persons described in the information sheet,
- had enough time to think about whether I want my child take part in this study,
- agree to allow my child take part in this study.

Tick as appropriate

I agree for my (or my child's) samples collected in this study to be used for this and future research that is approved by an appropriate ethics committee in The Gambia. This includes human genetic testing. Yes ☐ No☐

I agree for me/my child's samples to be shipped outside The Gambia following ethical approval as above Yes ☐ No☐

Participant's signature/ thumbprint* for **assent** (child aged 12-17 years) _____ _____
Signature if participant is an adult? Date (dd/mmm/yyyy) Time (24hr)

Participant's parent/guardian signature/thumbprint* _____ _____
Date (dd/mmm/yyyy) Time (24hr)

Printed name of witness* _____

Printed Name of Person obtaining consent _____

I attest that I have explained the study information accurately in _____ to, and was understood to the best of my knowledge by, the participant/parent/guardian. He/she has freely given consent to participate *in the presence of the above named witness (where applicable).

Signature of Person obtaining consent _____ _____
Date (dd/mmm/yyyy) Time (24hr)

** Only required if the participant is unable to read or write.*

APPENDIX III: STUDY QUESTIONNAIRE

Protocols For Obtaining Core Phenotypes

1 Participant's ID:	
2 Has the participant consented?	◯ Yes ◯ No
3 Date participant consented?	
4 For which country are you completing this survey	◯ Ghana ◯ Mali ◯ Gambia ◯ Cameroon ◯ Ethiopia ◯ Madagascar ◯ Kenya
5 In what language is the questionnaire being administered?	
Town and GPS	
6 Date of Interview/Examination:	
7 Name of Town/District	
8 GPS coordinate Elevation	
9 LATITUDE	
10 Longitude	
Social Demography	
11 Is your date of birth known?	◯ Yes ◯ No
14 Is your Age known	◯ Yes ◯ No
16 Please indicate your age group	◯ 0 - 5 years ◯ 6 - 10 years ◯ 11 - 15 years ◯ 16 - 20 years ◯ 21 - 25 years ◯ 26 - 30 years ◯ 31 - 35 years ◯ 36 - 40 years ◯ 41 - 45 years ◯ 46 - 50 years
17 Is your mother's age known	◯ Yes ◯ No

18 Enter Mother's age	
19 Gender?	◯ Male ◯ Femal
20 What is the main language spoken by your Father?	◯ Mandinka ◯ Wollof ◯ Jola ◯ Fulani ◯ Serehule ◯ Serre ◯ Other
21 What is your ethnic or tribal affiliation?	◯ Mandinka ◯ Wollof ◯ Jola ◯ Fulani ◯ Serehule ◯ Serre ◯ Other
22 What is your Mother's ethnic or tribal affiliation?	◯ Mandinka ◯ Wollof ◯ Fulani ◯ Serehule ◯ Serre ◯ Other
malaria information	
23 Have you had Malaria before	◯ Yes ◯ No ◯ Don't Know
24 When?	◯ Currently ◯ In past one month ◯ In past two month ◯ Other
25 What was the type of Malaria diagnosed?	◯ P.ovale ◯ P.vivax ◯ P.malariae ◯ P.falciparum ◯ Don't Know
26 Is this a relapse	◯ Yes ◯ No ◯ Don't Know
27 Malaria treatment in the past month?	◯ Yes ◯ No
28 Do you use malaria control intervention?	◯ Yes ◯ No
29 List of control intervention	☐ Indoor Residual Spray ☐ Bednet use Seasonal Malaria Chemoprevention ☐ Other
30 How often do you use bednets	☐ Always ☐ Sometimes

Migrations (in and out of villages between seasons)

31 Have you made any overnight trip outside of your town or village in the last 3 months?

- ◯ Yes
- ◯ No
- ◯ Can't remember

Confirmation by study staff

32 Confirm By Study Staff

33 Confirm By - Staff Code

35 APP Username

36 REDCap Username

37 Confirmed Date

APPENDIX IV: RNA EXTRACTION PROTOCOL

- Before you Start, Please Clean Bench and Biosafety Cabinet with 10 % bleach, 70 % Ethanol and RNAse Away™ Reagent (Thermo Fisher Scientific).
- Always Wear Gloves When Handling RNA
- Keep a Box of Sterile 200 µl, 1000 µl, 10 µl, and 20 µl Tips in a Clean Biosafety Cabinet Just for RNA Work
- Keep a Bag of Sterile Eppendorfs Just for RNA Work

Lysis and homogenisation (Adapted from Qiazol, Trizol and TriReagent books

1) TriReagent/Trizol: to 200 µl of RBC at 50% Hc add 750 µl TriReagent

Mix thoroughly for 1-2 min (can vortex in pulses) then freeze at –80 C

Extraction and Purification

2) For each 1ml (Trizol + sample) add 200 µl chloroform, shake vigorously 15 sec by hand

3) Place at RT for 10 min

4) Spin 13 000 rpm for 15 min at 4 C
(3 phases: upper colourless aqueous phase with RNA; white interphase with DNA, lower red organic phase with protein)
Volume of upper phase is 60% of the Trizol used in step 1

5) During the 15 min spin prepare master mix for step 10 as follows: for each sample, mix 16.25 µl of DNAse I + 113.75 µl of RDD buffer. Make a master mix by multiplying these volumes by the number of tubes you have and a little bit extra (eg if you have 2 samples make enough for 2 and a half samples, therefore, you will need 40.62 µl of DNAse I (i.e 16.25 x 2.5) and 284.3 µl of RDD buffer (i.e 113.75 x 2.5). Mix by inverting NOT vortexing

6) When the 15 min spin has finished transfer the upper phase to a new eppendorf tube
Note: Qiazol/Trizol/TriReagent contain phenol. Follow adequate procedures for discharge

Adapted from Rneasy Micro Handbook
7) Add 1 volume of 70 % ethanol (eg 1 ml Trizol+sample, ≈ 600 µl ethanol) and mix vigorously (vortex)

8) Apply max 700 µl of the above to an Rneasy MinElute Spin Column in a 2ml tube, spin 15 sec at 10 000 rpm; discard the flow-through; repeat if necessary (i.e if you have a volume larger than 700 µl you will have to repeat this step until all the samples has been through the column)

9) Pipette 350 µl Buffer RW1 into the Rneasy MinElute Spin Column, incubate 5 min RT, spin 15 sec at 10 000 rpm; discard the flow-through; reuse collection tube

10) Pipette 80 µl of master mix (prepared in step 5) onto the silica-gel membrane of the column (without touching the membrane), incubate 15 min RT

11) Pipette 350 µl Buffer RW1 into the Rneasy Column, spin 15 sec at 10 000 rpm; discard the flow-through; reuse collection tube

12) Pipette 50 µl of master mix onto the silica-gel membrane of the column (without touching the membrane), incubate 15 min RT

13) Pipette 350 µl Buffer RW1 into the Rneasy Column, spin 15 sec at 10 000 rpm; discard the flow-through; reuse collection tube

14) Transfer the column into a new 2ml collection tube. Pipette 500 µl Buffer RPE (previously diluted with ethanol!) onto the column, spin 15 sec at 10 000 rpm; discard the flow-through; reuse collection tube

15) Add 500 µl of 80% ethanol to the column, spin 2 min at min 8000g (10 000 rpm); discard the flow-through and the collection tube

16) Transfer the column into a new 2ml collection tube. Open the cap of the spin column and spin 5 min at full speed; discard the flow-through and the collection tube. Note that when placing the tubes in the centrifuge there should be one space in between them and caps should face the opposite direction of rotation.

17) To elute transfer the spin column to a new 1.5 ml collection tube, pipette 16 µl RNAse-free water onto the silica-gel membrane without touching the membrane; spin 1min at 10 000 rpm to elute. Aliquot the RNA into two eppendorfs. Label eppendorfs with RNA, study number and date.

Quantification of RNA

Take 1 µl and dilute in 49 µl of water (i.e 1:50 dilution) in a disposable uvette. Read in Giorgio's spectophotometer. Multiply result by 50 to obtain µg/ml concentration.

Notes Before starting:
- Prepare 70 % ethanol (35 ml 96-100% ethanol plus 15 ml RNAse free water)
- Prepare 80 % ethanol (40 ml 96-100% ethanol plus 10 ml RNAse free water)
- Make sure RPE is diluted with ethanol
- To prepare Dnase I stock solution dissolve the 1500 Units in 550 µl of Rnase-free water. Note: spin down the tube with the powder before opening it to avoid losses. Mix by inverting not by vortexing. For long-term storage aliquot at – 20 C (up to 9 months). Keep thawed aliquots at 4 C for 6 weeks without refreezing

APPENDIX V: SUPPLEMENTARY DATA

Table S 1: MSP2 allele typing PCR constituents and cycling conditions

PCR Constituents and Conditions					
MSP2 Primary PCR (pPCR)			**MSP2 *nested* PCR (nPCR)**		
Reagent	**Final Concentration**	**Volume (µl) x 1**	**Reagent**	**Final Concentration**	**Volume (µl) x 1**
Thermopol buffer	1X	1.5	Thermopol buffer	1X	2
dNTPs	125 Um	0.19	dNTPs	100 µM	0.2
MSP2: S2-forward primer	250 nM	0.38	MSP2: (allelic family) forward	125 nM	0.25
MSP2: S3-reverse primer	250 nM	0.38	MSP2: stail rev primer	125 nM	0.25
Taq polymerase		0.076 unit	Taq polymerase	0.05 unit	0.2
DNA Template		2	MSP2 pPCR Product		1
Nuclease free water		10.35	Nuclease free water		16.1
Total		**15**			**20**
PCR Cycling conditions			**PCR cycling conditions**		
Steps	**Temperature (º C) and Time**	**Number of cycles**	**Steps**	**Temperature (º C) and Time**	**Number of cycles**
Initial Denaturation	94, 5 minutes	1	Initial Denaturation	95, 5 minutes	1
Denaturation	94, 30 seconds	25	Denaturation	95 30 seconds	30
Annealing	58, 2 minutes		Annealing	58, 1 minute	
Extension	72, 2 minutes		Extension	72, 1 minute	
Final extension	72, 5 minutes	1	Final extension	72, 5 minutes	1

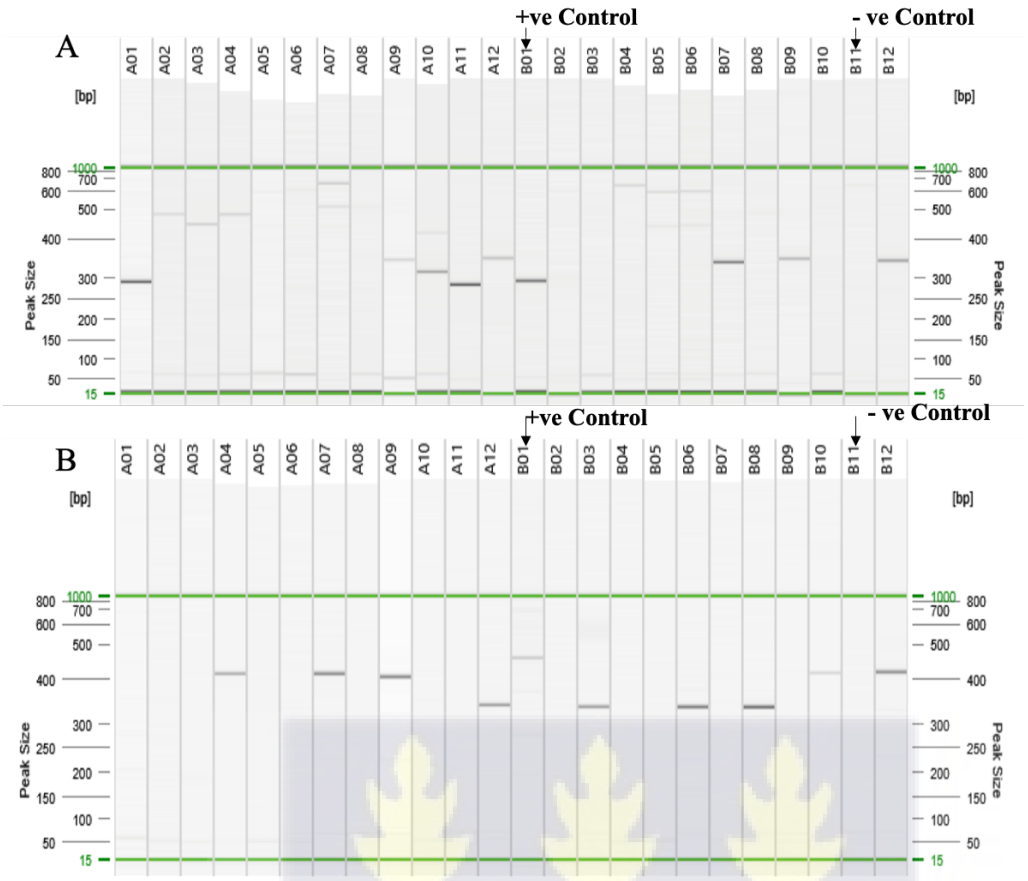


Figure S 1: QIAxcel electroporegram for MSP gene showing band sizes in base pair (bp).

A) 3D7 allelic family and B) FC27 allelic family aligned with a marker (15-1000)bp. A1-B12 are the genotyped samples for both 3D7 and FC27, bp = base pairs, +ve = positive control, and -ve =negative control.

Table S 2: Primers/probes sequences and cycling conditions for ligand gene analysed by reverse transcriptase quantitative PCR (RT-qPCR)

PCR Constituents and Conditions				
Gene name	**Primers and probes Sequences (5'-3')**	**Cycling conditions**	**Primer Reference**	**Polymerase**
Rh5	ACGAAGAATCAAGAAAATAATCTGACGTTACT	95°C for 10 min; 39 cycles of 95°C for 30 sec, 50°C for 1 min:30 sec, 60°C for 1 min	Nery et al., 2006	TaqMan Universal PCR Master Mix (Thermo Fisher Scientific; Cat4326708)
	TGTTGAATGATCTTTAGCATTATTTGTTTTTATATTCTCTTT			
	Fam-CTTCTTCAGTGCTCTTTATTG-BHQ-1			
Rh4	GAAATGACGCAATTCCCTCAAAAGA			
	GGTGTGTTTTATTTATATCATGTTGATTCTGTGA			
	Hex-CTTCTGGATCGTTTTTTTTT-BHQ-1			
Clag2	TGTAATCACACATAGAACATATGAAACAC			
	CTCTCCCTTATTTTCTATCTCTTTTGATTC			
	Tet-AGTCCAATTGAGAATCCTGGAGAGTTTTTTAGGAA-BHQ-1			
Rh2b	ACAGAAAGCGATGATATTGATAACAGTGAA			
	CCCATGGGTGTTACTTCTATGACT			
	Tet-CATATGAGTGACATCGAAAGTA-BHQ-1			
EBA-181	GCGGGTAGTACAATATTAGATGATTC		(Blair, 2002)	
	TGTTGTGTGCTAAAA-TTATGTTCTTG			
	Fam-AAATGACAGAAGGTAGCGAAAGTGATGTTGGAG-TAMRA			
EBA-175	AATTTCTGAAAAATATTGTGACCATATG	95°C for 10 min; 39 cycles of 95°C for 30 sec, 56°C for 1 min:30 sec, 60°C for 1 min	Blair et al., 2002	
	ATGAAGAAATCCCATTAAAAACATGCACTAAAGA			
	Hex-GATACTGCACAACACAGATTTCTTG-BHQ-1			
AMA1 (House keeping gene)	GAAATGACGCAATTCCCTCAAAAGA		Nery et al., 2006	
	GGTGTGTTTTATTTATATCATGTTGATTCTGTGA			
	Fam-CTTCTGGATCGTTTTTTTTT-MGB			

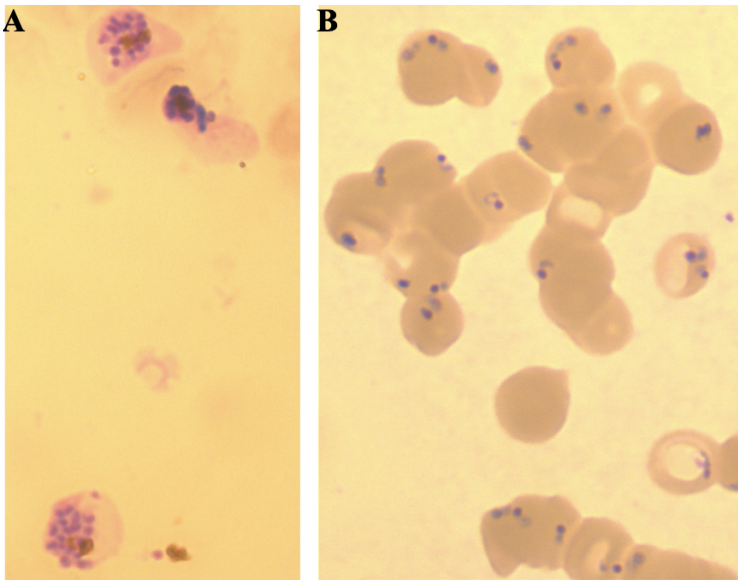


Figure S 2: Microscopy images of *P. falciparum* clinical isolates. (A) Mature hyper-segmented schizont stage parasites before freezing in Trizol for RNA extraction. B) Ring stage parasites before invasion assay set up

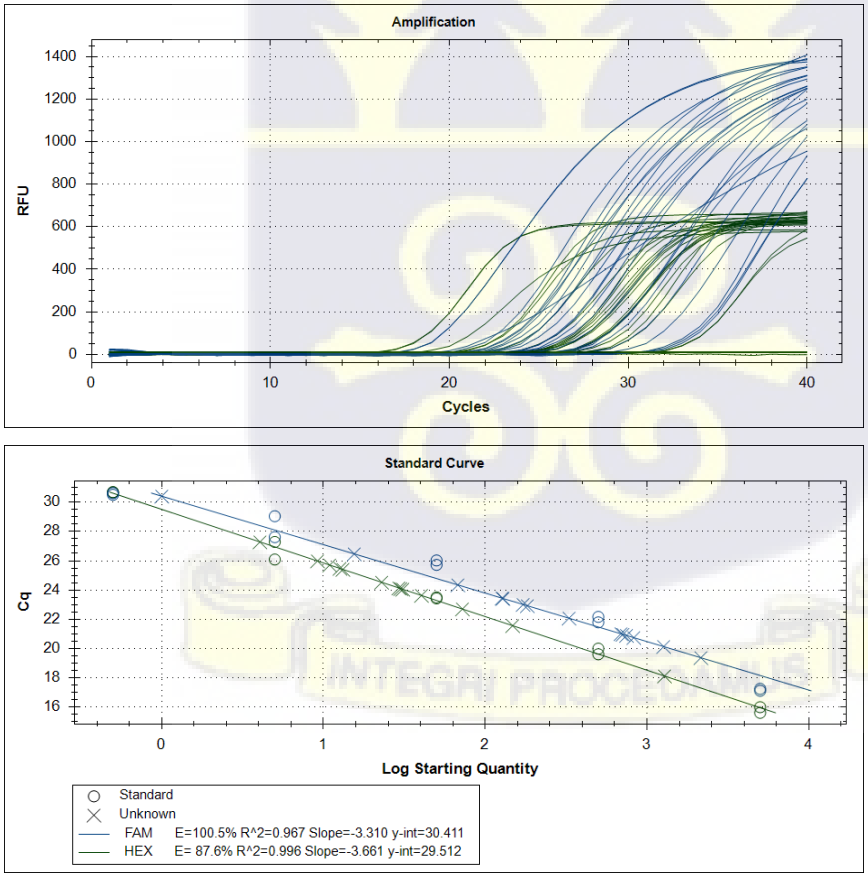


Figure S 3: RT-qPCR amplification and standard curves

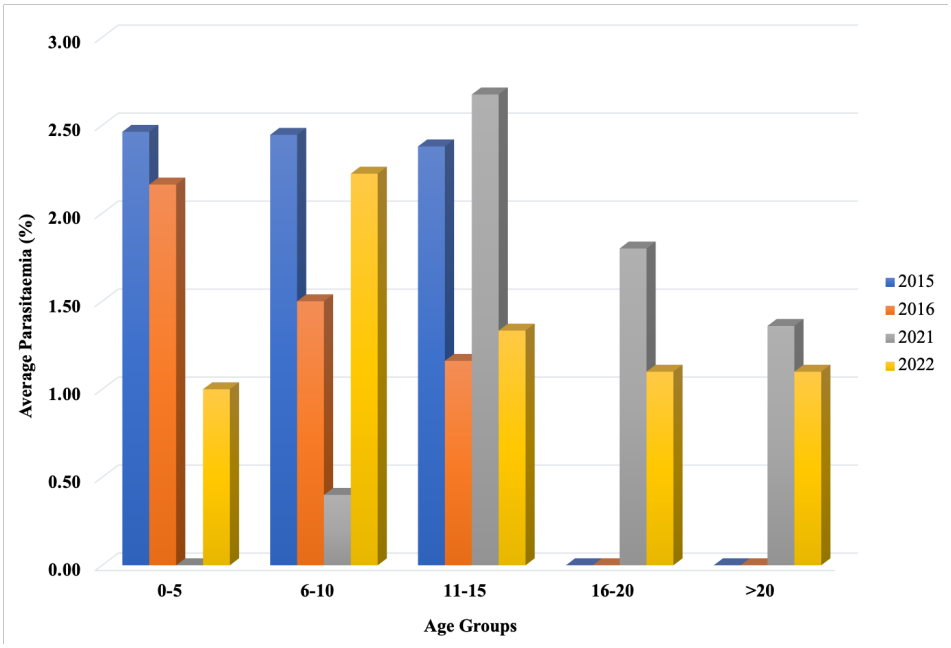


Figure S 4: Average parasitaemia of study participants according to age groups and sampling years

Table S 3: Genes, primers, and PCR conditions for amplification of *P. falciparum* invasion ligand genes and erythrocyte receptors

***P. falciparum* Targets**						
Gene name	**Primer Sequence (5'-3')**	**Targeted Region/exon**	**Product Size**	**Primer Reference**	**Cycling conditions**	**Polymerase**
Erythrocyte binding antigen 140 (EBA140)	CTGAAATATCTATTGGAAAGG GAATTATGGCACATTCTTTACT	region II	526	(Ochola-Oyier et al., 2016)	10 cycles (94°C for 5 mins, 94°C for 15 sec, 55.8°C for 30 sec, 72°C for 2 mins) and 25 cycles (94°C for 15 sec, 55.8°C for 30 sec, 72°C for 2 mins +5 sec/cycle), final elongation 72°C for 7 min	Taq DNA Polymerase
Erythrocyte binding antigen 181 (EBA181)	TTGTAGATTAATGAAGGAGA GTTACATTTTTGTCTTTAGCA	region II	1196	(Ochola-Oyier et al., 2016)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 51°C for 30 sec, 68°C for 3 min; final extension step of 72°C for 5 min	Taq DNA Polymerase
Erythrocyte binding antigen 175 (EBA175)	GGAAGAAATACTTCATCTAATAACG CATCCTTTACTTCTGGACACATCG	region II	1810	(Baum, Thomas, et al., 2003)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 51°C for 30 sec, 68°C for 3 min; final extension step of 72°C for 5 min	Qiagen Fast Cycling master mix
Erythrocyte binding like protein 1 (EBL1)	GTACCCCTGAATATAAAGTTCCTTTTTT GGTTATCTGTACAAATCCATTTCATTTC)	N'terminal region	663	(Githui et al., 2010)	10 cycles (94°C for 5 mins, 94°C for 15 sec, 55.8°C for 30 sec, 72°C for 2 mins) and 25 cycles (94°C for 15 sec, 55.8°C for 30 sec, 72°C for 2 mins +5 sec/cycle), final elongation 72°C for 7 min	Taq DNA Polymerase
Cytoadherence linked asexual protein 2 (CLAG2)	TATATGGAAAAAGTAGTAATATACAGG TACTAGTATGTGGTTGAT ATTCTTTTG		702	(Alexandre et al., 2011)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 53°C for 30 sec, 68°C for 2 min; final extension step of 72°C for 5 min	Taq DNA Polymerase

Cytoadherence linked asexual protein 8 (CLAG8)	GTTTATGGAAAAAGTGGTAAAATAGG CTCTTTAAGTTTTCTTCTGAATAGTTC		750	(Alexandre et al., 2011)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 53°C for 30 sec, 68°C for 2 min; final extension step of 72°C for 5 min	Qiagen Fast Cycling master mix
Merozoite Surface Protein 6 (MSP6)	TCAAAATGAATAAGATTTATAATATTAC CTAAATAGATGGATCATTTCTT	full length	1285	(Ochola-Oyier et al., 2016)	10 cycles (94°C for 5 mins, 94°C for 15 sec, 55.8°C for 30 sec, 72°C for 2 mins) and 25 cycles (94°C for 15 sec, 55.8°C for 30 sec, 72°C for 2 mins +5 sec/cycle), final elongation 72°C for 7 min	Taq DNA Polymerase
Duffy binding-like merozoite surface protein (DBLMSP)	TGATTCGAATCTAAGAAACG GAAATAAATCTGTCATATCTTC	full length	1993	(Ochola et al., 2010)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 50°C for 30 sec, 72°C for 2 min; final extension step of 72°C for 5 min	Taq DNA Polymerase
Suface-associated interspersed prottein 4.2 (SURFIN 4.2)	TCCCATTTTTGATAATATGC CACACTCTTTTCAATAATTACT		3228	(Ochola et al., 2010)	94°C for 5 mins x 1 cycle-(94°C for 15 sec, 53.6°C for 30 sec, 68°C for 4 mins) x 10 cycles – (94°C for 15 sec, 53.6°C for 30 sec, 68°C for 4 min + 5 sec/cycle)x25 cycles, 72°C for 7 min x1 cycle	Qiagen Fast Cycling master mix
Reticulocyte binding protein homologue 4 (Rh4)	TCAAATAATGCTGTTAATGC GTAAAATTTAGGTATGTATGTTTG	CR1 binding region	842	(Ochola-Oyier et al., 2019)	10 cycles (94°C for 5 mins, 94°C for 15 sec, 50°C for 30 sec, 72°C for 2 mins) and 25 cycles (94°C for 15 sec, 50°C for 30 sec, 72°C for 2 mins +5 sec/cycle), final elongation 72°C for 7 min	Taq DNA Polymerase
Reticulocyte binding protein homologue (Rh5)	CGA AGA ATC AAG AAA ATA ATC TG GAATATTCATTTGACATGTCT	exon 2	1500	(Ochola-Oyier et al., 2019)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 51°C for 30 sec, 72°C for 2 min; final extension step of 72°C for 5 min	Taq DNA Polymerase

Merozoite Surface Protein 1 (MSP1)	CATTGAGACCTTATACAATAAC TTAGAGGAACTGCAGAAAATACCA	33kDa region and 19kDa region	600	(Ochola-Oyier et al., 2016; Takala et al., 2006)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 47°C for 30 sec, 72°C for 2 min; final extension step of 72°C for 5 min	Taq DNA Polymerase
Erythrocyte Receptor Targets						
Glycophorin A (GPA)	AGGCCAATAATACAATACTTACCA GACAGATTTATATTTAGAGGTTCC	2, 3, 4, 5	4124	(Ochola-Oyier et al., 2019)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 62°C for 30 sec, 68°C for 3 min; final extension step of 72°C for 5 min	Q5 Hot start high fidelity master mix
Glycophorin B (GPB)	CAATAATACAATACTTACCACGCA CATGAGACTTCATGTTATCTTGGA	2, 3, 4, 5	5541	(Ochola-Oyier et al., 2019)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 62°C for 30 sec, 68°C for 3 min; final extension step of 72°C for 5 min	Q5 Hot start high fidelity master mix
Glycophorin C (GPC)	GCATACTGCAGAGAACTTAAATG TACATACATAGATACGTACGATGTA	2, 3	4167	(Ochola-Oyier et al., 2019)	95°C for 5 min; 34 cycles of 95°C for 15 sec, 62°C for 30 sec, 68°C for 3 min; final extension step of 72°C for 5 min	Q5 Hot start high fidelity master mix
Complement receptor 1 (CR1)	GTGATAGATAGTCCTTTGAT TGAAGGACAGATTGCACAGAA GTTTAGTGACTCATGAGATTTC CAAATACTAATCTCCTGATCCAAC	4, 5	324, 547	(Ochola-Oyier et al., 2019)	10 cycles (94°C for 5 mins, 94°C for 15 sec, 58°C for 30 sec, 72°C for 2 mins) and 25 cycles (94°C for 15 sec, 58°C for 30 sec, 72°C for 2 mins +5 sec/cycle), final elongation 72°C for 7 min	Qiagen Fast Cycling master mix

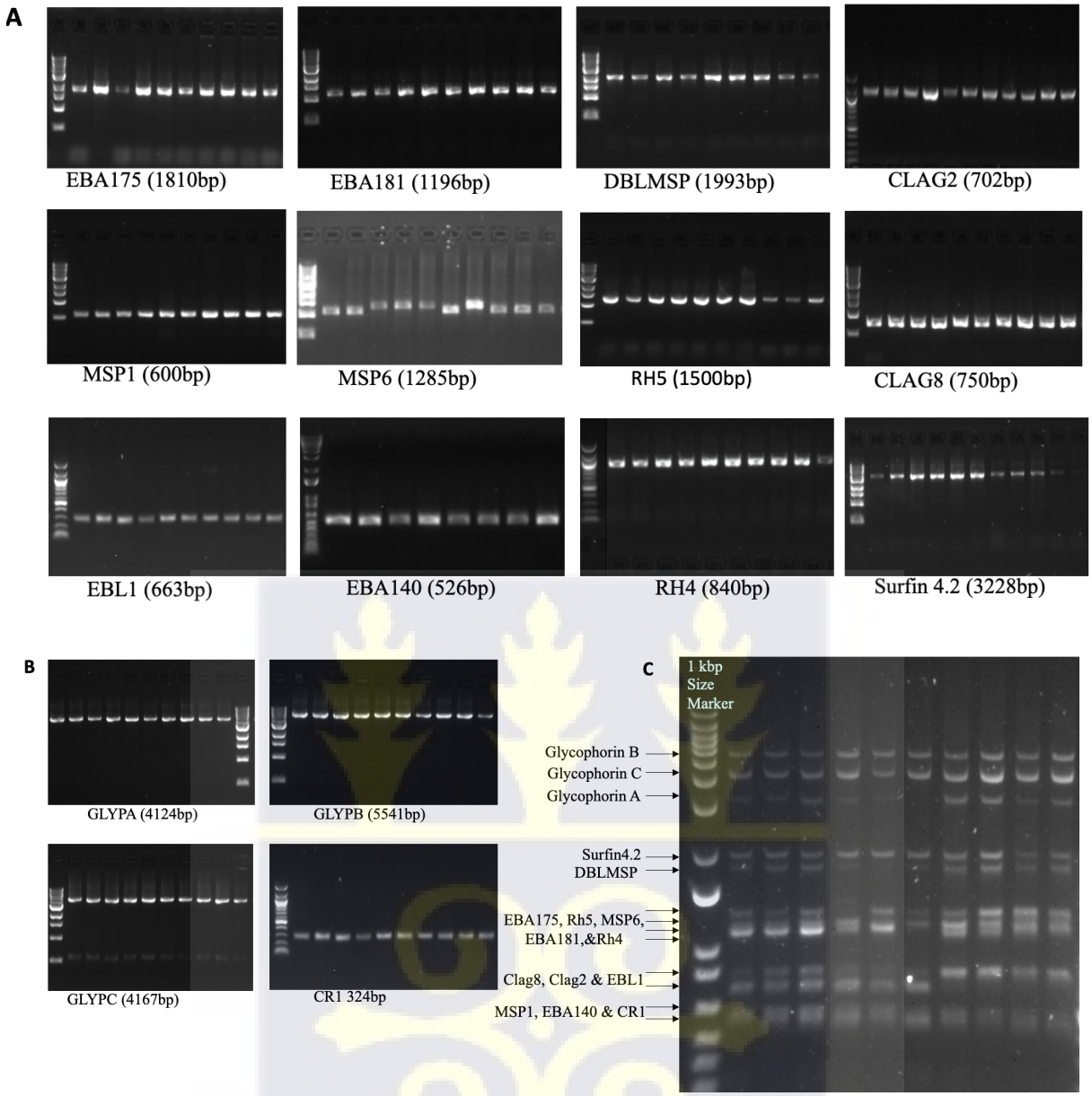


Figure S 5: Gel images of parasite ligands (12) and human receptors (4). (A) parasite ligands, B) human receptors, and C) pooled amplicons before library preparation. The fragment size of each target is indicated in the bracket

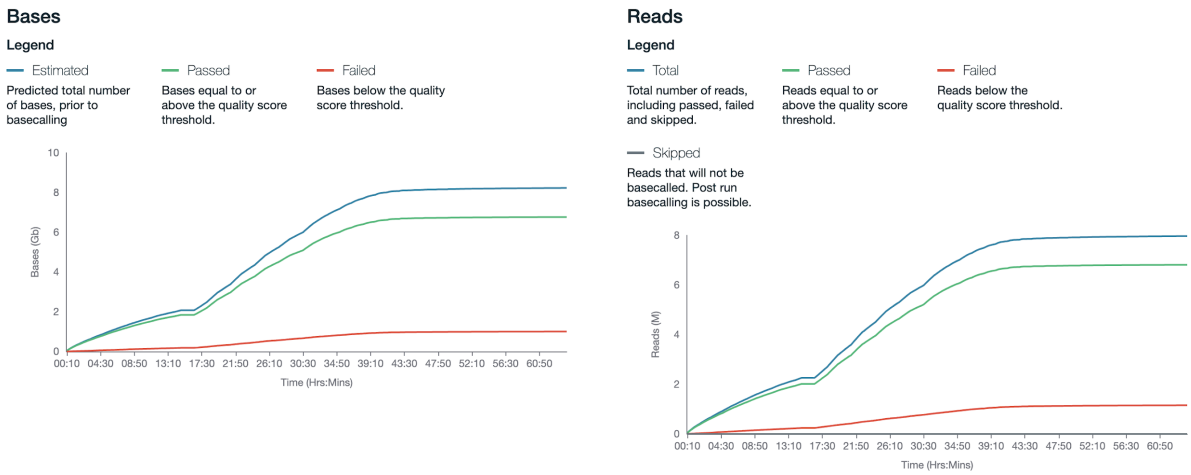


Figure S 6: Quality control of nanopore data in real-time; cumulative output showing the total amount of bases and reads sequenced over time

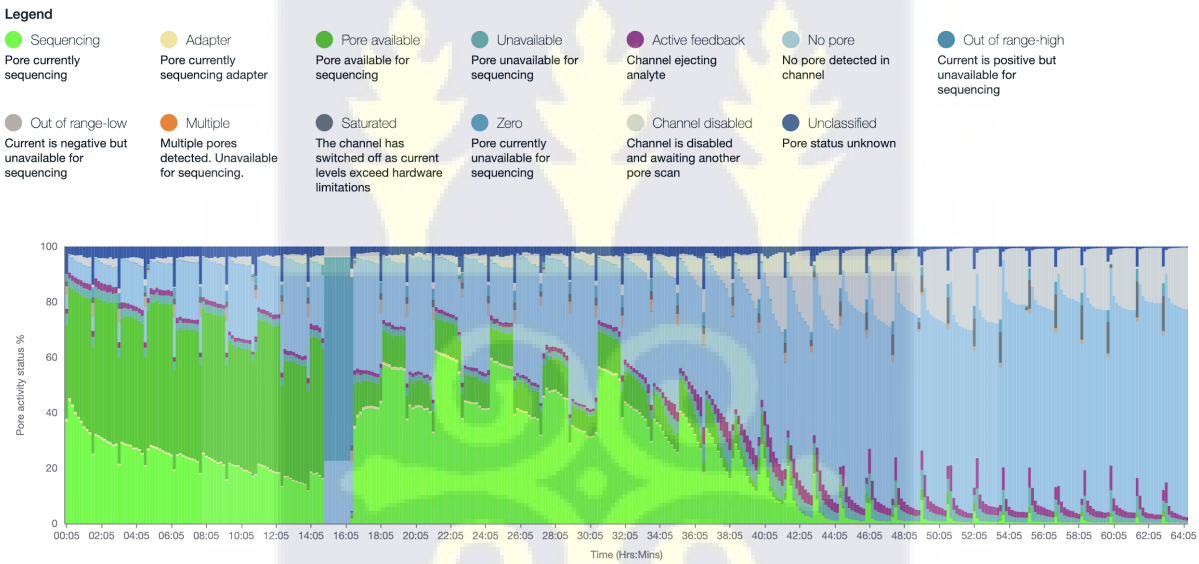


Figure S 7: Quality control of nanopore data in real-time: pore activity graph showing samples performance during sequencing

Table S 4: Distribution of ethnicities of sampling population per site

Ethnicity	Sites				
	Basse (%)	Brikama (%)	Fajikunda (%)	EFSTH (%)	Total
Fulani	22 (31.9)	6 (6.6)	5 (7.1)	8 (13.8)	41
Jahanka	1 (1.5)	0 (0)	0 (0)	0 (0)	1
Jola	0 (0)	3 (3.3)	3 (4.3)	3 (5.2)	9
Mandinka	24 (34.8)	37 (40.7)	27 (38.6)	13 (22.4)	101
Serehule	8 (11.6)	0 (0)	3 (4.3)	2 (3.5)	13
Wolof	1 (1.5)	3 (3.3)	27 (38.6)	25 (43.1)	56
Kiriyol	0 (0)	1 (1.1)	0 (0)	0 (0)	1
NA	13 (18.8)	41 (45.1)	5 (7.1)	7 (12.1)	66
Total	69	91	70	58	**288**

Table S 5: Univariable logistic regression model table indicating significant SNPs (Highlighted) associated with severe malaria

Gene	SNP_Position	OR	LL_95CI	UL_95CI	P_value
CR1	207590351	1.754	0.354	8.701	0.491
	207590415	2.885	0.477	17.454	0.249
	207590415	2.961	0.603	14.544	0.181
	207590435	1.088	0.372	3.182	0.877
	207590447	0.793	0.302	2.083	0.638
	207590600	3.132	0.666	14.728	0.148
	207590613	2.540	0.536	12.043	0.240
	207590614	1.691	0.394	7.257	0.480
	207590614	2.336	0.675	8.089	0.181
	207590648	1.476	0.133	16.390	0.751
	207590648	1.648	0.606	4.484	0.328
	207590660	0.776	0.305	1.970	0.593
	207590745	1.170	0.458	2.992	0.743
	207590793	**5.489**	**1.182**	**25.489**	**0.030**
	207590807	**4.228**	**1.142**	**15.656**	**0.031**
	207590810	3.302	0.900	12.113	0.072
	207590845	6.032	0.758	48.023	0.090
	207591627	1.330	0.463	3.817	0.596
	207591868	0.515	0.169	1.572	0.244
	207591876	0.985	0.169	5.734	0.986
	207591876	**0.324**	**0.104**	**1.006**	**0.051**
	207591892	0.621	0.216	1.784	0.376

	207591919	**3.400**	**1.133**	**10.203**	**0.029**
	207591924	1.126	0.394	3.217	0.825
	207591945	1.228	0.468	3.223	0.677
	207592005	1.254	0.470	3.344	0.651
	207592214	0.341	0.064	1.817	0.207
	207592214	1.143	0.311	4.201	0.841
	207592218	3.000	0.459	19.592	0.251
	207592218	2.550	0.528	12.314	0.244
	207592246	0.202	0.023	1.809	0.153
	207592246	0.593	0.215	1.630	0.311
	207592271	2.320	0.776	6.937	0.132
	207592275	5.068	0.627	40.982	0.128
	207592323	**4.750**	**1.598**	**14.118**	**0.005**
	207592371	3.259	0.691	15.376	0.135
	207632588	**6.250**	**1.344**	**29.068**	**0.019**
	207632593	**5.588**	**1.504**	**20.767**	**0.010**
	207632720	**4.378**	**1.350**	**14.199**	**0.014**
GPC	**126691057**	**4.667**	**0.984**	**22.138**	**0.052**
	126692091	4.348	0.924	20.455	0.063
	126692544	0.125	0.013	1.245	0.076
	126692544	0.404	0.061	2.657	0.346
GPA	144014115	0.444	0.032	6.188	0.546
	144014115	1.115	0.117	10.643	0.925
	144016829	0.583	0.093	3.653	0.565
	144016829	1.089	0.203	5.854	0.921
	144016855	0.294	0.042	2.046	0.216
	144016855	0.694	0.122	3.944	0.681
	144018106	0.231	0.016	3.371	0.284
	144018106	1.200	0.116	12.411	0.878
	144133195	1.484	0.163	13.494	0.726
CLAG2	pf3d7_02_v3843193	0.500	0.105	2.388	0.385
	pf3d7_02_v3843193	0.605	0.162	2.259	0.455
SURFIN4.2	pf3d7_04_v31099887	3.194	0.362	28.176	0.296
	pf3d7_04_v31099887	1.714	0.190	15.469	0.631
	pf3d7_04_v31100164	2.500	0.492	12.703	0.269
	pf3d7_04_v31100164	1.419	0.264	7.629	0.683
	pf3d7_04_v31100169	2.200	0.243	19.896	0.483
	pf3d7_04_v31100169	1.441	0.152	13.642	0.750
	pf3d7_04_v31100200	3.034	0.339	27.137	0.321
	pf3d7_04_v31100200	2.694	0.316	22.929	0.364
	pf3d7_04_v31100256	1.667	0.173	16.023	0.658
	pf3d7_04_v31100256	0.965	0.103	9.081	0.975

	pf3d7_04_v31100319	1.151	0.217	6.110	0.869
	pf3d7_04_v31100319	1.167	0.204	6.668	0.862
	pf3d7_04_v31100674	0.940	0.220	4.022	0.934
	pf3d7_04_v31100674	1.111	0.249	4.953	0.890
	pf3d7_04_v31100700	2.889	0.562	14.843	0.204
	pf3d7_04_v31100700	1.806	0.302	10.796	0.517
	pf3d7_04_v31100704	1.440	0.336	6.170	0.623
	pf3d7_04_v31100704	1.146	0.226	5.814	0.870
	pf3d7_04_v31100719	2.053	0.489	8.619	0.326
	pf3d7_04_v31100719	0.933	0.180	4.838	0.935
	pf3d7_04_v31100904	4.200	0.484	36.462	0.193
	pf3d7_04_v31100904	5.091	0.603	42.995	0.135
	pf3d7_04_v31100926	0.825	0.190	3.578	0.797
	pf3d7_04_v31100926	0.882	0.171	4.565	0.881
	pf3d7_04_v31100990	5.871	0.698	49.384	0.103
	pf3d7_04_v31100990	2.500	0.264	23.669	0.424
	pf3d7_04_v31101249	0.485	0.117	2.005	0.318
	pf3d7_04_v31101249	0.838	0.263	2.666	0.765
	pf3d7_04_v31101252	0.643	0.150	2.761	0.552
	pf3d7_04_v31101252	0.643	0.201	2.052	0.456
	pf3d7_04_v31101254	1.026	0.369	2.849	0.961
	pf3d7_04_v31101257	0.579	0.223	1.504	0.262
	pf3d7_04_v31101272	3.000	0.406	22.176	0.282
	pf3d7_04_v31101272	1.333	0.406	4.379	0.635
	pf3d7_04_v31101284	1.880	0.182	19.450	0.596
	pf3d7_04_v31101284	**3.463**	**1.145**	**10.476**	**0.028**
	pf3d7_04_v31101287	1.040	0.173	6.258	0.966
	pf3d7_04_v31101287	1.470	0.471	4.586	0.507
	pf3d7_04_v31101294	1.560	0.313	7.777	0.587
	pf3d7_04_v31101294	1.058	0.338	3.309	0.923
	pf3d7_04_v31101301	2.359	0.559	9.963	0.243
	pf3d7_04_v31101301	1.721	0.438	6.769	0.437
	pf3d7_04_v31101304	1.077	0.237	4.902	0.924
	pf3d7_04_v31101304	0.800	0.193	3.314	0.758
	pf3d7_04_v31101309	1.500	0.330	6.822	0.600
	pf3d7_04_v31101309	1.320	0.330	5.276	0.695
	pf3d7_04_v31101341	1.594	0.607	4.190	0.344
	pf3d7_04_v31101342	1.038	0.410	2.630	0.937
	pf3d7_04_v31101344	1.033	0.102	10.496	0.978
	pf3d7_04_v31101344	1.138	0.394	3.288	0.811
	pf3d7_04_v31101346	2.318	0.193	27.885	0.508
	pf3d7_04_v31101346	1.030	0.377	2.817	0.954

	pf3d7_04_v31101348	1.688	0.613	4.649	0.312
	pf3d7_04_v31101362	0.737	0.294	1.847	0.515
	pf3d7_04_v31101367	0.571	0.071	4.593	0.599
	pf3d7_04_v31101367	1.655	0.336	8.165	0.536
	pf3d7_04_v31101370	1.820	0.330	10.047	0.492
	pf3d7_04_v31101370	1.219	0.234	6.347	0.814
	pf3d7_04_v31101373	1.000	0.158	6.330	1.000
	pf3d7_04_v31101373	1.313	0.397	4.347	0.655
	pf3d7_04_v31101387	0.933	0.092	9.507	0.954
	pf3d7_04_v31101387	1.358	0.453	4.067	0.585
	pf3d7_04_v31101395	0.836	0.140	5.011	0.845
	pf3d7_04_v31101395	1.115	0.367	3.386	0.847
	pf3d7_04_v31101407	0.262	0.029	2.388	0.235
	pf3d7_04_v31101407	0.762	0.277	2.094	0.598
	pf3d7_04_v31101414	0.821	0.078	8.604	0.870
	pf3d7_04_v31101414	1.396	0.426	4.575	0.581
	pf3d7_04_v31101418	1.044	0.397	2.745	0.931
	pf3d7_04_v31101420	0.857	0.313	2.344	0.764
	pf3d7_04_v31101424	0.688	0.153	3.087	0.625
	pf3d7_04_v31101424	0.504	0.179	1.417	0.194
	pf3d7_04_v31101442	0.353	0.055	2.247	0.270
	pf3d7_04_v31101442	0.686	0.195	2.406	0.556
	pf3d7_04_v31101449	1.500	0.220	10.218	0.679
	pf3d7_04_v31101449	1.910	0.398	9.163	0.418
	pf3d7_04_v31101453	0.554	0.093	3.312	0.517
	pf3d7_04_v31101453	0.794	0.255	2.477	0.691
	pf3d7_04_v31101462	0.700	0.113	4.329	0.701
	pf3d7_04_v31101462	1.250	0.373	4.185	0.717
	pf3d7_04_v31101466	0.737	0.129	4.210	0.731
	pf3d7_04_v31101466	1.131	0.290	4.414	0.859
	pf3d7_04_v31101492	5.838	0.703	48.454	0.102
	pf3d7_04_v31101492	5.400	0.592	49.256	0.135
	pf3d7_04_v31101578	1.188	0.293	4.820	0.810
	pf3d7_04_v31101578	1.118	0.321	3.894	0.861
	pf3d7_04_v31101587	1.375	0.326	5.796	0.664
	pf3d7_04_v31101587	0.444	0.087	2.276	0.330
	pf3d7_04_v31101648	0.931	0.207	4.196	0.926
	pf3d7_04_v31101648	0.612	0.140	2.671	0.514
	pf3d7_04_v31101652	1.231	0.318	4.758	0.763
	pf3d7_04_v31101652	0.909	0.249	3.313	0.885
	pf3d7_04_v31101664	1.132	0.370	3.464	0.829
	pf3d7_04_v31101664	0.660	0.236	1.848	0.429

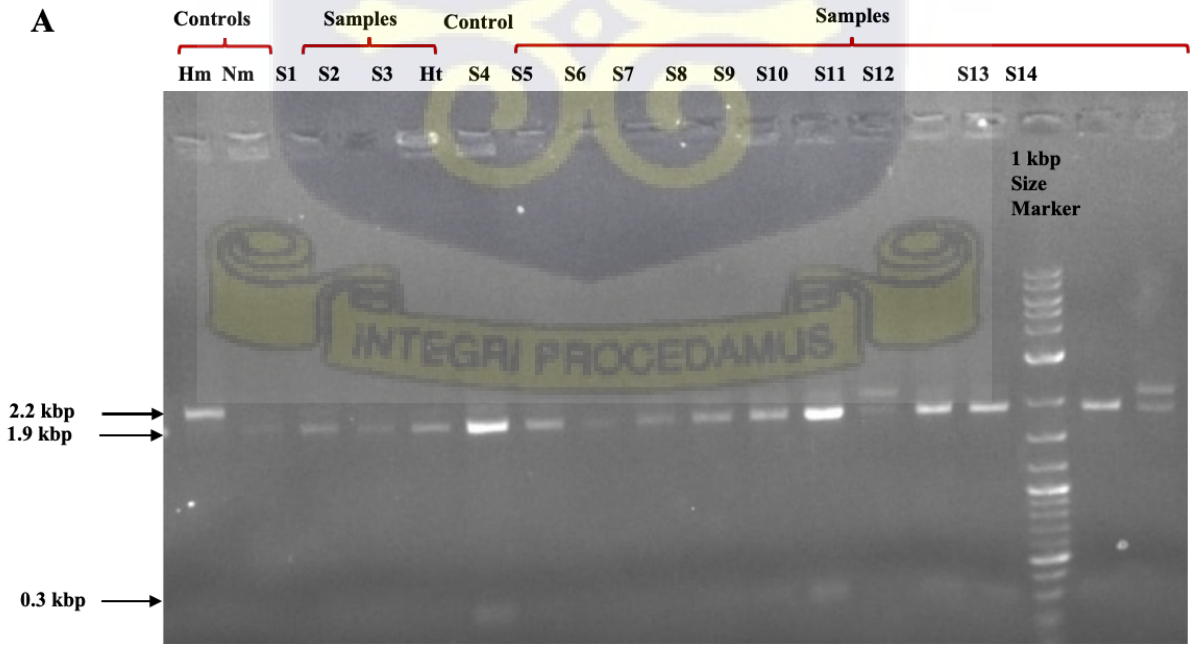
	pf3d7_04_v31101672	0.621	0.068	5.715	0.674
	pf3d7_04_v31101672	1.017	0.331	3.120	0.977
	pf3d7_04_v31101673	1.190	0.107	13.299	0.887
	pf3d7_04_v31101673	2.778	0.744	10.371	0.128
	pf3d7_04_v31101676	0.785	0.241	2.556	0.688
	pf3d7_04_v31101681	0.915	0.360	2.325	0.852
	pf3d7_04_v31101688	1.100	0.108	11.155	0.936
	pf3d7_04_v31101688	2.031	0.679	6.069	0.205
	pf3d7_04_v31101696	0.479	0.046	5.031	0.540
	pf3d7_04_v31101696	2.509	0.673	9.351	0.171
	pf3d7_04_v31101707	2.545	0.257	25.173	0.424
	pf3d7_04_v31101707	4.175	0.511	34.091	0.182
	pf3d7_04_v31101709	1.000	0.127	7.848	1.000
	pf3d7_04_v31101709	3.109	0.659	14.663	0.152
	pf3d7_04_v31101725	1.875	0.154	22.880	0.622
	pf3d7_04_v31101725	4.412	0.549	35.481	0.163
	pf3d7_04_v31101734	0.717	0.104	4.933	0.736
	pf3d7_04_v31101734	1.608	0.328	7.891	0.559
	pf3d7_04_v31101735	0.842	0.196	3.615	0.817
	pf3d7_04_v31101735	0.933	0.299	2.917	0.906
	pf3d7_04_v31101739	1.667	0.165	16.821	0.665
	pf3d7_04_v31101739	2.769	0.332	23.082	0.346
	pf3d7_04_v31102061	1.750	0.318	9.644	0.520
	pf3d7_04_v31102061	1.784	0.360	8.854	0.479
	pf3d7_04_v31102363	1.250	0.373	4.193	0.718
	pf3d7_04_v31102363	0.533	0.127	2.232	0.389
	pf3d7_04_v31102393	2.049	0.229	18.339	0.521
	pf3d7_04_v31102393	1.960	0.205	18.721	0.559
EBA175	pf3d7_07_v31358875	1.733	0.272	11.054	0.561
	pf3d7_07_v31358875	2.758	0.549	13.860	0.218
MSP6	pf3d7_10_v31414245	4.875	0.519	45.789	0.166
	pf3d7_10_v31414245	3.391	0.403	28.562	0.261
	pf3d7_10_v31414254	4.231	0.427	41.873	0.217
	pf3d7_10_v31414254	2.918	0.345	24.717	0.326
	pf3d7_10_v31414266	5.500	0.568	53.215	0.141
	pf3d7_10_v31414266	2.694	0.316	22.948	0.365
	pf3d7_10_v31414268	4.615	0.469	45.390	0.190
	pf3d7_10_v31414268	3.184	0.378	26.781	0.287

Table S 6: Primers, PCR constituents, and conditions for glycophorin B variants

PCR Constituents and Conditions						
	Deletion 1			**Deletion 2**		
Components	**Primer sequence**	**Volume per reaction (µl)**	**Final concentration**	**Primer sequence**	**Volume per reaction (µl)**	**Final concentration**
Forward Primer	GGACTGCCG CATGTTCAG	0.25	250 nM	GGTCATGAGAAAACG TTTGAATTTTCCAG	0.75	750 nM
Reverse Primer	CTCTGGTAG CCCTCCTCAA G	0.25	250 nM	CAGTTCTGCCAACTCT CATCTT	0.25	250 nM
Qiagen Fast Cycling master mix (x2)		2.5	0.5x		2.5	0.5x
Water		6			5.5	
DNA 20 ng/µl		1	2 ng/µl		1	2 ng/µl
Total		10			10	
	Deletion 1 PCR Cycling Conditions			**Deletion 2 PCR Cycling Conditions**		
	Step	**Condition**	**Time**	**Step**	**Condition**	**Time**
	Initial denaturation	95 °C	5 min	Initial denaturation	95 °C	5 min
	Denaturation	96 °C	5 sec	Denaturation	96 °C	5 sec
	Annealing	63 °C	6 sec	Annealing	59 °C	6 sec
	Elongation	68 °C	3 min 15 sec	Elongation		3 min 15 sec
	Go to Step	2 x 39 times		Go to Step	2 x 39 times	
	Final elongation	72 °C	1 min	Final elongation	72 °C	1 min
	Hold	15 °C	Infinity	Hold	15 °C	Infinity

Table S 7: Glycophorin B variants restriction digest constituents and conditions

Deletion 1			**Deletion 2**		
Components	**Volume per reaction**	**Final concentration**	**Components**	**Volume per reaction**	**Final concentration**
AciI (10U/ µl)	0.3	0.3U/ µl	BsrBI (10U/ µl)	0.3	0.3U/ µl
CutSmart (10X)	1	1X	CutSmart (10X)	1	1X
Water	3.7		Water	3.7	
PCR Product	5		PCR Product	5	
Incubation time: 37 ºC for 9 hours, 65 ºC for 15 mins			**Incubation time**: 37 ºC for 5 hours, 65 ºC for 15 mins		
	Band sizes per genotype				
	Deletion 1		**Deletion 2**		
	Wild type	1.9 kb and 0.3 kb	Wild type	2.2 kb	
	Heterozygote	2.2 kb, 1.9 kb, and 0.3 kb	Heterozygote	2.2 kb, 1.3 kb, and 0.9 kb	
	Homozygote	2.2 kb	Homozygote	1.3 kb and 0.9 kb	
	1 % agarose Gel				
	Run time: 1 hour at 100 V				



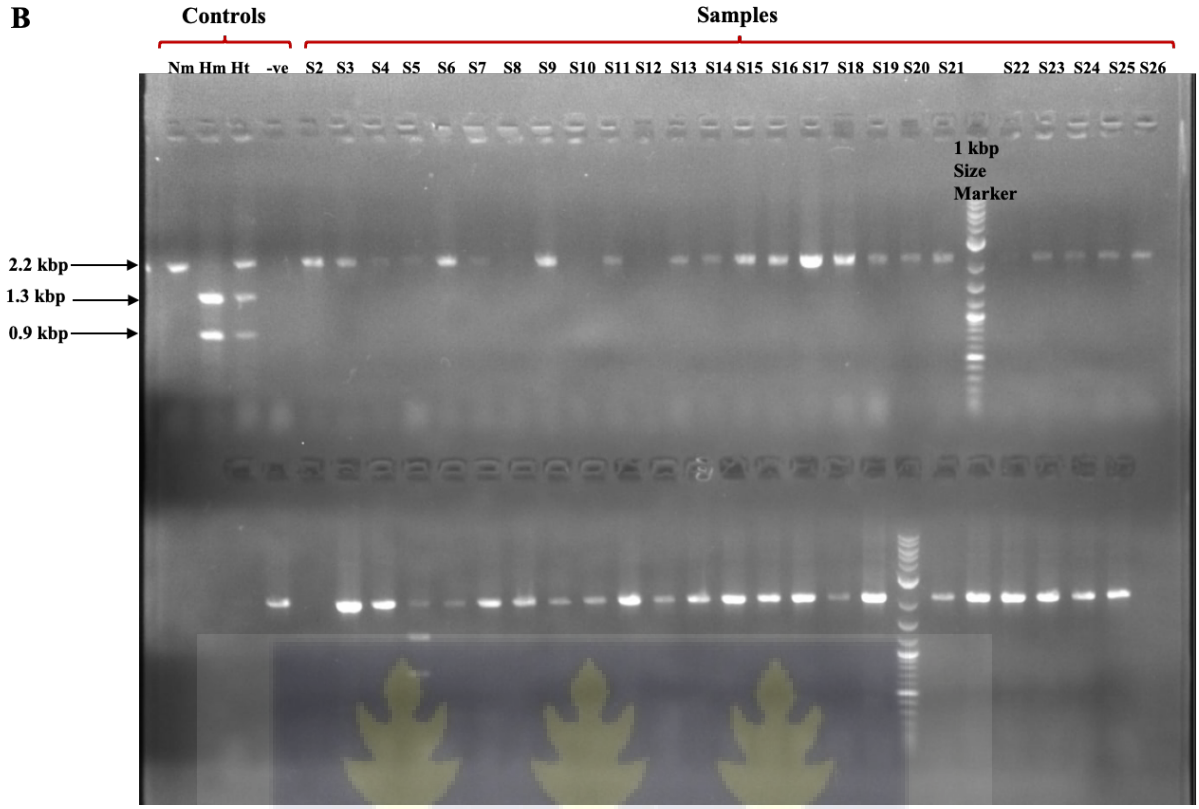


Figure S 8: Gel images of GPB deletions by RFLP. A) DEL 1 digested PCR product: Acil enzyme cuts the non-DEL 1 but does not cut DEL 1 Alleles (homozygous is left uncut), B) DEL 2 digested PCR product: BsrBI enzyme cuts the DEL 2 alleles but does not cut non-DEL 2 alleles. Nm = normal or wild type, Hm = homozygous, Ht = heterozygous, kbp = kilobase pairs, and -ve = negative control.

Table S 8: Glycophorin B deletions multiplex PCR Constituents, conditions and expected fragment sizes

Multiplex PCR Constituents			
Components	**Primer sequence**	**Volume per reaction (µl)**	**Final concentration**
Del-B-Left-F	GAGCAATGGTCTGATATCAA GAAAGGC	0.2	200 nM
Del-B-Left-R	TGAGATTCTGCCCCAGGCTA CCT	0.2	200 nM
Del-B-Left-WT-R	GGGTCTGTGGGTTCTCTCAG TG	0.2	200 nM

Del-B-Right-F	TTTGGACTCCATGTCTCACA TCCAGT	0.2	200 nM
Del-B-Right-R	GATTTAGGTTCCTTGAAGGT AATGACTACA	0.2	200 nM
Del-B-Right-WT-F	CTGCTTTCACGGGCTGTTAT CCAA	0.25	250 nM
Qiagen Fast Cycling master mix (x2)		2.5	0.5x
Water		4.25	
DNA		2	
Total		10	
PCR Cycling Conditions			
	Step	**Condition**	**Time**
	Initial denaturation	95 ºC	5 min
	Denaturation	95 ºC	30 secs
	Annealing	60 ºC	30 secs
	Elongation	72 ºC	3 min
	Go to Step	2 x 34 times	
	Final elongation	72 ºC	5 min
	Hold	4 ºC	infinity
Expected fragment sizes			
Variants	**Genotypes**	**Fragment sizes**	
DEL 1	Homozygous (L/L)	**1004; 553** (DEL 1 Left, WT right)	
DEL 1	Heterozygous (L/WT)	**1004; 553; 507** (DEL 1 left, WT right, WT left)	
DEL 2	Homozygous (R/R)	**671; 507** (WT left, DEL 2 right)	
DEL 2	Heterozygous (R/WT)	**671; 553; 507** (DEL 2 right, WT right, WT left)	
DEL 1 and DEL 2	Compound Heterozygous (L/R)	**1004; 671; 553; 507** (DEL 1 left, DEL 2 right, WT left, WT right)	
Non-DEL 1/DEL 2	Wild type (WT/WT)	**553; 507** (WT right, WT left)	

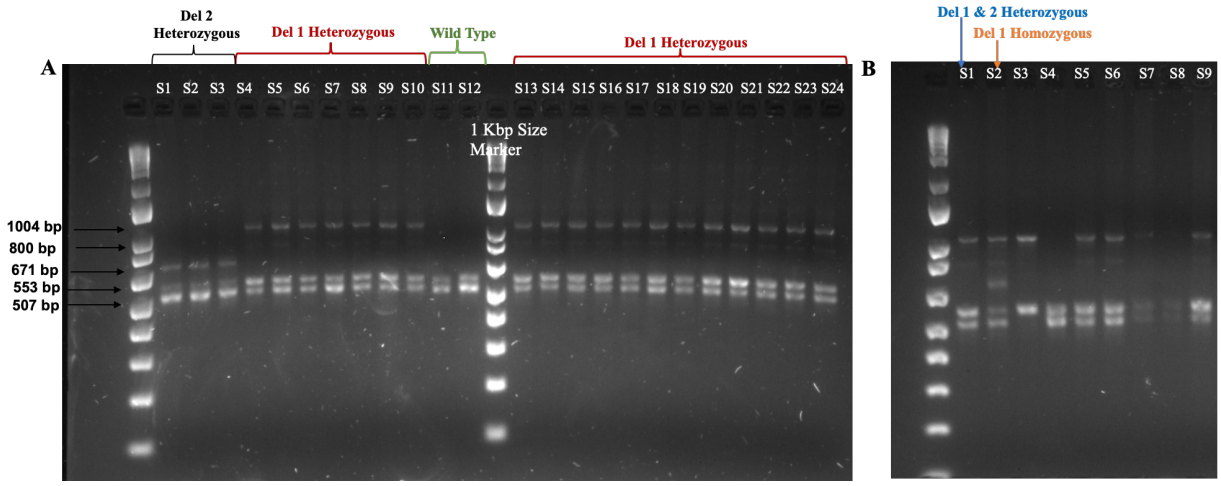


Figure S 9: Gel image of GPB DEL 1 and DEL 2 using multiplex PCR approach. DEL 2 heterozygous with three band sizes (507, 553, and 671bp), DEL 1 heterozygous with three band sizes (507, 553, and 1004bp), and wild type with two band sizes (507, and 553), B) DEL 1 and 2 heterozygous or compound heterozygous with four band sizes (507, 553, 671, and 1004), and DEL 1 homozygous with two band sizes (507 and 671).

Table S 9: HbSS PCR constituents, conditions, and expected fragment sizes

Gene	**Primer sequence**	**Fragment Sizes**	
HbSS Codon 6 (rs334)	CGGCTGTCATCACTTAGACCT	HbSS: 382bp	
	AGGGTGGGAAAATAGACCAA	HbAS: 382bp, 202bp, 180bp	
		HbAA: 202bp, 180bp	
PCR Cycling Conditions			
	Step	**Condition**	**Time**
	Initial denaturation	95 ºC	5 minutes
	Denaturation	95 ºC	30 secs
	Annealing	60 ºC	30 secs
	Elongation	72 ºC	30 secs
	Go to Step	2 x 34 times	
	Final elongation	72 ºC	5 minutes
	Hold	4 ºC	infinity
Restriction Digest Conditions			
Components	**Volume per reaction**	**Incubation time**	
DdeI (10U/ µl)	1 µl	37 C for 2 hours, 65 oC for 20 mins	
CutSmart (10X)	2 µl		
Water	4 µl		
PCR Product	3 µl		

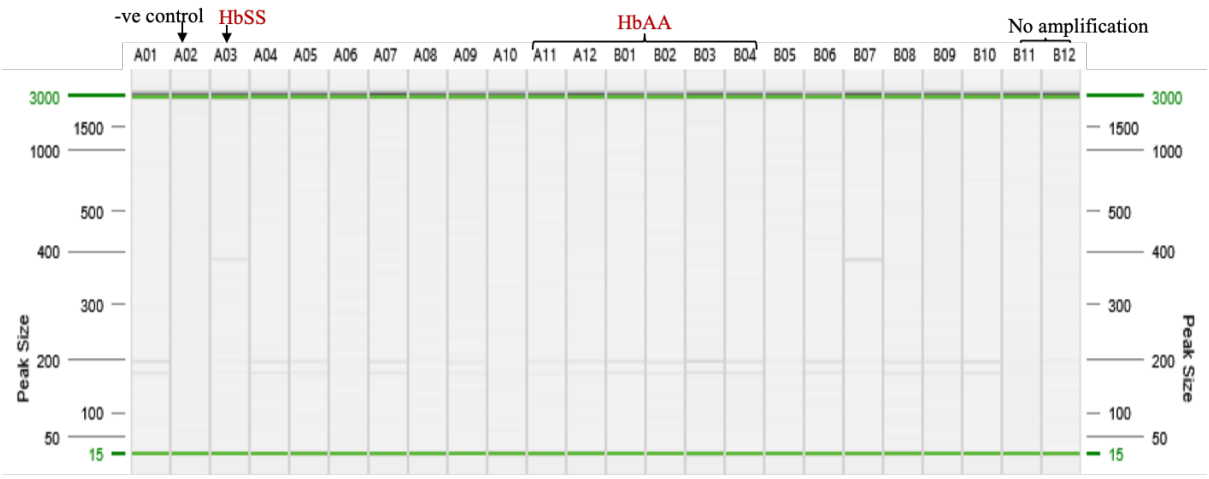


Figure S 10: QIAxcel electroporegram for HbSS showing band sizes in base pair (bp).