GENETIC FACTORS
IMPLICATED IN ASYMPTOMATIC MALARIA INFECTIONS

JOSEPH GISAINA AMWOMA

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN GENETICS OF THE UNIVERSITY OF EMBU

AUGUST, 2024

DECLARATION

This thesis is my original work and has not been presented elsewhere for a degree or any other award.

Signature………………………………… Date………………………………….

Gisaina Joseph Amwoma

B520/1220/2018

Department of Biological Sciences.

This thesis has been submitted for examination with our approval as the University Supervisors.

Signature………………………………… Date………………………………….

Dr. Sarah N. Kituyi

Department of Biological Sciences

University of Embu.

Signature…………………………………. Date………………………………….

Dr. Ben Andagalu

Malaria Drug Resistance Laboratory

United States Army Medical Research Directorate/KEMRI.

Signature…………………………………. Date………………………………….

Dr. Franklin N. Nyabuga

Department of Biological Sciences

University of Embu.

DEDICATION

I dedicate this work to my family members: my father, mother, and my siblings. Thank you for your encouragement and support.

ACKNOWLEDGEMENT

I extent my gratitude to Dr. Ben Andagalu for giving me the opportunity to carry out this research. I greatly thank my university supervisors, Dr. Sarah N. Kituyi and Dr. Franklin N. Nyabuga, for their academic and moral support, guidance, enthusiasm as well as encouragement throughout the work. I am thankful to Mrs. Gladys Chemwor for the outstanding assistance she offered at different levels of the project.

I acknowledge members of the USAMRD-K MDR laboratory for their great support and Mr. Benjamin Opot for his support in sequencing, library preparation and bioinformatics. Ms. Martha Kivecu for the support in bioinformatics analyses. I also acknowledge the Global Emerging Infections Surveillance (GEIS) program for funding this project.

Thanks to my parents and siblings for the moral and financial support and encouragement offered during this period. Above all, I thank almighty God for helping me make it and for the protection he has offered me throughout the research period.

TABLE OF CONTENTS

DECLARATION ii
DEDICATION iii
ACKNOWLEDGEMENT iv
LIST OF TABLES viii
LIST OF FIGURES ix
LIST OF APPENDICES x
LIST OF ABBREVIATIONS /ACRONYMS xi
ABSTRACT xii
CHAPTER ONE 1
INTRODUCTION 1
  1.1 Background information 1
  1.2 Statement of the Problem 2
  1.3 Justification 3
  1.4 Research Questions 3
  1.5 Research Objectives 3
    1.5.1 Main Objective 3
    1.5.2 Specific Objectives 4
CHAPTER TWO 5
LITERATURE REVIEW 5
  2.1 The Malaria Burden 5
  2.2 Malaria Parasites 6
  2.3 Genotyping of *Plasmodium* Parasites 7
  2.4 Selected Malaria Outcomes 8
    2.4.1 Uncomplicated Malaria 8
    2.4.2 Asymptomatic Malaria 9
  2.5 *Plasmodium falciparum* Virulent Genes and Surface Antigens 10
    2.5.1 *Plasmodium falciparum var* Gene Expression 11
    2.5.2 Single Nucleotide Polymorphisms Within *P. falciparum var* Genes 11
  2.6 Malaria Transcriptomics 12
    2.6.1 Epigenetic Regulation in Malaria 12
    2.6.2 Non-coding RNAs 13
  2.7 Host-parasite Interactions 14

CHAPTER THREE 15
MATERIALS AND METHODS 15
3.1 Study Site 15
3.2 Ethical consideration 15
3.3 Study Design and Population 15
3.3.1 Uncomplicated Malaria Participants 15
3.3.2 Asymptomatic Participants 17
3.4 Sampling Criteria and Sample Size 18
3.5 Retrieval of Malaria Transcriptome Data from GEO 18
3.6 Sample Collection and Storage 18
3.7 Sample Processing and Sequencing 19
3.7.1 RNA Extraction from Blood Samples 19
3.7.2 Ethanol Precipitation of RNA 19
3.7.3 RNA Quantification and Integrity Analysis 20
3.7.4 Molecular Detection of *Plasmodium* Species 20
3.7.5 *Plasmodium* Speciation Assay 21
3.7.6 Genomic DNA Digestion and mRNA Enrichment 22
3.7.7 Library Preparation 23
3.8 Transcription Patterns of *P. falciparum* Parasites During Asymptomatic Malaria Infections 24
3.8.1 Quality Assessment of the Raw Sequence Data 24
3.8.2 Gene Expression Analysis 24
3.8.3 Gene ontology and Pathway Analysis 24
3.8.4 Long noncoding RNA Characterization 25
3.8.5 Analysis of Differentially Expressed Long noncoding RNA 25
3.9 The Expression of the *var* Genes Subsets in ASM versus UM Carriers 25
3.9.1 Extraction of *PfEMP1* Encoding and Regulator Genes 25
3.9.2 *Plasmodium falciparum* Virulence Genes Expression Analysis 26
3.10 Single Nucleotide Polymorphisms Within *var* Genes Associated with Asymptomatic Malaria 26
3.10.1 Alignment to the pf3D7 Reference Genome 26
CHAPTER FOUR 27
RESULTS 27
4.1 Demographic Characteristics of Study Participants 27

4.2 *Plasmodium* Detection and Speciation of Malaria Samples............................ 27
4.3 RNA Sequencing.. 28
4.3.1 RNA Quantification and Integrity Analysis 28
4.3.2 Library Preparation.. 29
4.4 Characteristics of the RNA Sequence Datasets... 30
4.4.1 Quality Assessment of the Raw Sequence Dataset.................................. 30
4.4.2 Dataset Retrieval from NCBI 31
4.4.3 Alignment and Transcript Assembly.. 31
4.5 Transcription Patterns of *P. falciparum* Parasites During Asymptomatic Malaria Infections Compared to Symptomatic Infections..................................... 32
4.5.1 Comparative Analysis of DEGs Between ASM Carriers Versus UM....... 32
4.5.2 Comparative Analysis of DEGs Between ASM Carriers Versus Naïve Malaria Infections 36
4.5.3 Function Classification by Gene Ontology of Differentially Expressed Genes in both ASM vs. UM and ASM and NM.. 39
4.5.4 Functional Enrichment of the DEGs in ASM vs. UM and ASM vs. NM. 44
4.5.5 Differential Expression of Long Non-coding RNA in Asymptomatic Malaria Compared to Uncomplicated Malaria 47
4.5.6 Analysis of the Interaction of lncRNAs with Host Immune Genes........... 48
4.6 Expression of the *var* Genes Subsets in Asymptomatic Malaria Compared to Uncomplicated Malaria 49
4.6.1 Expression of *var* genes subsets in asymptomatic malaria compared to symptomatic malaria.. 49
4.7 Single Nucleotide Polymorphisms within *var* Genes Associated with Asymptomatic Malaria 50
CHAPTER FIVE.. 53
DISCUSSIONS, CONCLUSIONS AND RECOMMENDATIONS......................... 53
5.1 Discussions.. 53
5.2 Conclusions 57
5.3 Limitations... 57
5.4 Recommendations 58
REFERENCES... 59

LIST OF TABLES

Table 3.1 Primer and Probe Sequences for the qualitative real PCR Reaction 20
Table 3.2. Consensus primers for Plasmodium speciation assay 22
Table 4.1: Gametocyte carriage in both asymptomatic and uncomplicated malaria samples. 28
Table 4.2: A representative of the quantity and the quality results of RNA isolated from both asymptomatic and uncomplicated malaria samples. 29
Table 4.4: Characteristics of the raw sequence dataset 31
Table 4.5: Alignment Summary Statistics 32
Table 4.6: Top 10 upregulated genes in asymptomatic compared to uncomplicated malaria. 35
Table 4.7: Downregulated genes in asymptomatic versus uncomplicated malaria infections. 36
Table 4.8: Top 10 upregulated genes in asymptomatic versus naïve malaria infections 39
Table 4.9: Top 10 Enriched molecular function terms in asymptomatic vs. uncomplicated malaria. 40
Table 4.10. Representative of enriched molecular function terms in asymptomatic vs. uncomplicated malaria. 41
Table 4.11: Top 10 enriched cellular component gene ontology terms in asymptomatic vs. naïve malaria. 42
Table 4.12: List of genes implicated in evasion of host immune response 45
Table 4.13: The interaction of upregulated long noncoding RNAs with P. falciparum genes implicated in immune evasion. 49
Table 4.14: A table showing the 20 SNPs identified through variant calling in the P. falciparum virulent genes. 51
Table 4.15: The distribution of the 20 single nucleotide polymorphisms across the chromosomes 52

LIST OF FIGURES

Figure 2.1. Illustration of the full life cycle of Plasmodium parasites. 7
Figure 2.2. The structure of PfEMP1 proteins and the organization of the DBL domains. 11
Figure 4.1. Principal component analysis demonstrating the variation and clustering observed among samples from individuals with ASM compared to those with UM. 33
Figure 4.2. A volcano plot showing differentially expressed genes in ASM vs. UM. 34
Figure 4.3. Illustration of a principal component analysis demonstrating the variation and clustering observed among samples from asymptomatic individuals versus naïve malaria individuals. 37
Figure 4.4. A volcano plot Illustrating differentially expressed genes in ASM vs NM 38
Figure 4.5. Scatterplots depicting enriched Gene Ontology terms for asymptomatic compared to uncomplicated malaria, and asymptomatic versus naïve malaria infections.. 43
Figure 4.6 Scatter plot illustrating the top 14 Enriched KEGG Pathways 46
Figure 4.7. Cluster analysis of 23 (ASM) vs 12 (UM) samples by PCA. 47
Figure 4.8. Hierarchical clustering of differentially expressed lncRNA 48
Figure 4.9. MA plot showing the significantly enriched genes in asymptomatic malaria compared to uncomplicated malaria 50
Figure 4.10. A bar plot showing the number of effects by region 52

LIST OF APPENDICES

Appendix 1. Ethical Approval 70

Appendix 2. Characteristics of the filtered reads. 76

Appendix 3. KEGG pathways associated with the Differentially expressed genes in ASM vs. UM 78

Appendix 4. single nucleotide polymorphisms Identified in *P. falciparum* virulent genes.. 79

LIST OF ABBREVIATIONS /ACRONYMS

CIDR	cysteine-rich interdomain region
DBL	Duffy binding like proteins
dNTP’s	Deoxynucleotide triphosphates
iRBCs	infected red blood cells
LAMP	Loop-mediated isothermal amplification
CDC	Centre for Disease Control
MOH	Ministry of Health
MSP	Merozoite surface protein
NTS	N- Terminal Sequence
MDR	Malaria Drug Resistance
*P*LDH	*Plasmodium* lactate dehydrogenase
PCR	Polymerase chain reaction
PfEMP1	*Plasmodium falciparum* erythrocyte membrane protein 1
*Pf*HRP	*Plasmodium* Histidine-Rich Protein
qRT-PCR	Reverse Transcription Quantitative Polymerase Chain Reaction
RDT	Rapid diagnostic tests
rRNA	Ribosomal ribonucleic acid
WHO	World Health Organization
ASM	Asymptomatic malaria
UM	Uncomplicated malaria
NM	Naïve symptomatic malaria
ncRNA	Non coding RNA
lncRNA	Long non coding RNA
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
GEO	Gene expression omnibus
GEIS	Global Emerging Infections Surveillance
USAMRD	US Army Medical Research Directorate

ABSTRACT

Plasmodium falciparum malaria infections can be classified as either asymptomatic or symptomatic which includes uncomplicated malaria, severe malaria, cerebral malaria, and naïve malaria. In symptomatic cases, individuals infected with the parasites display symptoms associated with malaria. In contrast, asymptomatic cases do not show any classical symptoms and, therefore, act as reservoirs of the parasite. The persistence of *P. falciparum* parasites in an asymptomatic malaria individual plays a significant role in maintaining transmission. However, the molecular mechanisms underlying *P. falciparum*'s persistence in asymptomatic infections remain largely unknown. This study hypothesised that changes in the parasite's gene expression during asymptomatic infections might enhance their adaptability and fitness and therefore contribute to their persistence in the host. To investigate this hypothesis, a transcriptome analysis was done to identify *P. falciparum*'s genetic factors that are associated with asymptomatic infections. Whole blood RNA profiles from 25 field samples (15 asymptomatic (ASM) and ten uncomplicated malaria (UM) samples) were analyzed. In addition, publicly available transcriptome datasets from whole blood samples of 35 individuals with asymptomatic malaria (ASM, n=11), uncomplicated malaria (UM, n = 12), and naïve-malaria (NM, n = 12) were included for differential gene expression analysis, long noncoding RNAs expression analysis, and single nucleotide polymorphisms analysis within the virulent genes of *P. falciparum.* Data analysis revealed 755 differentially expressed genes (DEGs) between asymptomatic carriers and individuals with uncomplicated malaria, and 1773 DEGs between NM and ASM. Among the top differentially expressed genes were numerous genes coding for proteins of unknown functions (PUFs). The ontology analysis of DEGs revealed activated pathways linked to host-parasite interactions, including antigenic variation, immune evasion, crystalloid, apoplast, and binding processes and supressed pathways associated with carbon metabolism. Additionally, the analysis identified 28 immune evasion genes associated with host-parasite and symbiotic interactions, such as cell adhesion, evasion of the host immune response, modulation by symbionts of host cellular processes, and responses to biotic and external biotic stimuli. Moreover, 237 differentially expressed noncoding RNAs were detected when comparing asymptomatic carriers (ASM) to those with uncomplicated malaria (UM). Among these, five RNAs were observed to interact with six immune evasion genes, indicating a potential role in modulating the immune response during asymptomatic infections. The analysis of *var* gene family expression did not reveal statistically significant differences in the expression levels of the *var* groups. However, two genes, CUFF.75 and CUFF.203, were upregulated in asymptomatic infections compared to uncomplicated malaria infections. These findings demonstrate that *P. falciparum* establishes asymptomatic infections by suppressing the central carbon metabolism and expressing immune evasion genes, which mediate sequestration to avoid clearance and adapt to the host's defenses, resulting in long-lasting chronic infections. Additionally, the study identifies potential biomarkers for detection of asymptomatic malaria. These findings provide novel insights into *P. falciparum* genetic factors that confer a fitness advantage during asymptomatic infections.

CHAPTER ONE

INTRODUCTION

1.1 Background information

Malaria continues to be a major global health issue, impacting numerous tropical and subtropical nations. As of 2022, there were 249 million documented cases and 608,000 fatalities, predominantly concentrated in sub-Saharan Africa (Venkatesan, 2024). *Plasmodium falciparum* is the predominant cause of malaria in the sub Saharan region and is characterized by both symptomatic and asymptomatic infections (Mace et al., 2018). Understanding the significant role of asymptomatic malaria in sustaining transmission is crucial for the World Health Organization's (WHO) goal of reducing malaria cases and deaths by 90% by 2030 (WHO, 2019), necessitating the understanding and accurate detection of these infections, which was the focus of this study.

Individuals acquiring *P. falciparum* species infection from a mosquito bite are prone to display symptoms consistent with classical case presentation for malaria within 8 to 14 days of the bite termed as symptomatic malaria (Bartoloni & Zammarchi, 2012; Lindblade et al., 2013). Symptomatic malaria infections can include Uncomplicated malaria, severe malaria cerebral malaria, and naïve malaria. While the period of onset of symptoms may vary, studies have noted that some individuals with detectable *P. falciparum* species infection delay showing any symptoms for an unforeseeable duration, hence referred to as asymptomatic (Mackintosh et al., 2004; Chen et al., 2016). Notably, while individuals with symptomatic infections readily present at the health facilities and get treated to deter transmission, the asymptomatic individuals, despite accounting for a large percentage of malaria infections do not visit health facilities and hence remain obscure to treatment and control interventions (Mooney et al., 2022). Thus, they maintain active transmission of malaria infections in a local setting (Cheaveau et al., 2019; Andagalu et al., 2023).

Harbouring *Plasmodium* parasites, especially *P. falciparum,* without portraying symptoms has been attributed to acquired immunity after several exposures to malaria in-

fections, anti-parasite immunity, and a factor of balance between a pro and anti-inflammatory response that appears to be inherent in parasite modulation of the host immune response (Frimpong et al., 2020). Whilst the host genetic factors have been largely studied, there is limited data on transcriptome, functional genomics, proteome, and phenotypic description of the parasites during asymptomatic infections. The understanding of *Plasmodium* parasite biology primarily stems from clinical malaria cases and lab-cultivated clones, which may not accurately reflect the characteristics of asymptomatic infections (Nyarko & Claessens, 2021).

The persistence of the parasites without host immune system clearance suggests a multifaceted relationship involving both host and parasite factors. These interactions include molecular mechanisms such as adhesive interactions of infected erythrocytes that govern parasite invasion, activation of specific host immune responses and epigenetic regulation (Xia et al., 2018; Zhu et al., 2010). These molecular mechanisms further facilitate *Plasmodium*'s adaptation to changing environments, particularly through epigenetic regulation. Epigenetics give rise to new phenotypes without altering the DNA makeup, allowing the parasite to effectively adjust to the host conditions —a phenomenon referred to as phenotypic plasticity (Witmer et al., 2020).

Transcriptional variations and epigenetic regulation thus play a significant role in plasticity and could be central in establishing asymptomatic infections. However, the transcriptional variations and molecular mechanisms responsible for the parasites' persistence during asymptomatic infections are elusive.

1.2 Statement of the Problem

Plasmodium. falciparum infections constitute the largest proportion of malaria cases and mortality rates, particularly within Africa and significantly impact the economies of affected populations. The global spending on malaria control and elimination is estimated at $4.3 billion annually, yet these efforts are hindered by asymptomatic infections. These infections harbour potential infectious gametocytes and act as a reservoir, silently maintaining active malaria transmission locally. Despite their significance,

asymptomatic infections are poorly understood. The parasite genotypic factors implicated in the parasite persistence in the host without being cleared by the immune system is elusive. This study aimed at identifying the genetic factors involved in asymptomatic malaria through transcriptome analysis.

1.3 Justification

Accurate detection and quantification of asymptomatic infections is crucial for resource allocation and policy making. More importantly, detection and treating asymptomatic infections is vital for controlling and reducing malaria morbidity, which in turn can enhance the economic growth of affected countries. Therefore, the identification of parasite genetic factors involved in establishing asymptomatic infections provide valuable insights into the biology of these infections and assist in the discovery of potential biomarkers essential for developing rapid diagnostic tools, vaccines, as well as tailored therapeutic treatments for managing the disease at both individual and regional levels.

1.4 Research Questions

i. What is the transcription profile of *P. falciparum* parasites isolated from asymptomatic carriers?

ii. What is the expression profile of the *var* genes subgroups in asymptomatic malaria carriers compared to uncomplicated malaria?

iii. What single nucleotide polymorphisms within *var* genes are associated with asymptomatic malaria?

1.5 Research Objectives

1.5.1 Main Objective

To investigate *P. falciparum* genetic factors implicated in asymptomatic malaria infections.

1.5.2 Specific Objectives

i. To determine transcription patterns of *P. falciparum* parasites during asymptomatic malaria infections.

ii. To determine the expression of the *var* genes subgroups in asymptomatic malaria carriers.

iii. To determine single nucleotide polymorphisms within *var* genes associated with asymptomatic malaria.

CHAPTER TWO

LITERATURE REVIEW

2.1 The Malaria Burden

Malaria is preventable and treatable, yet its prevalence and impact are substantial, particularly in the tropical regions of the world. Africa represents 94% of the overall cases and fatalities associated with malaria (WHO, 2020). Malaria mortality rate declined by 63% between the year 2000 and 2019 (WHO, 2022). Despite the declining malaria burden, there is a high rate of changing epidemiology characterized by spatial and temporal variability of the parasites, thus posing new challenges to efforts geared at controlling malaria (Nkumama et al., 2017).

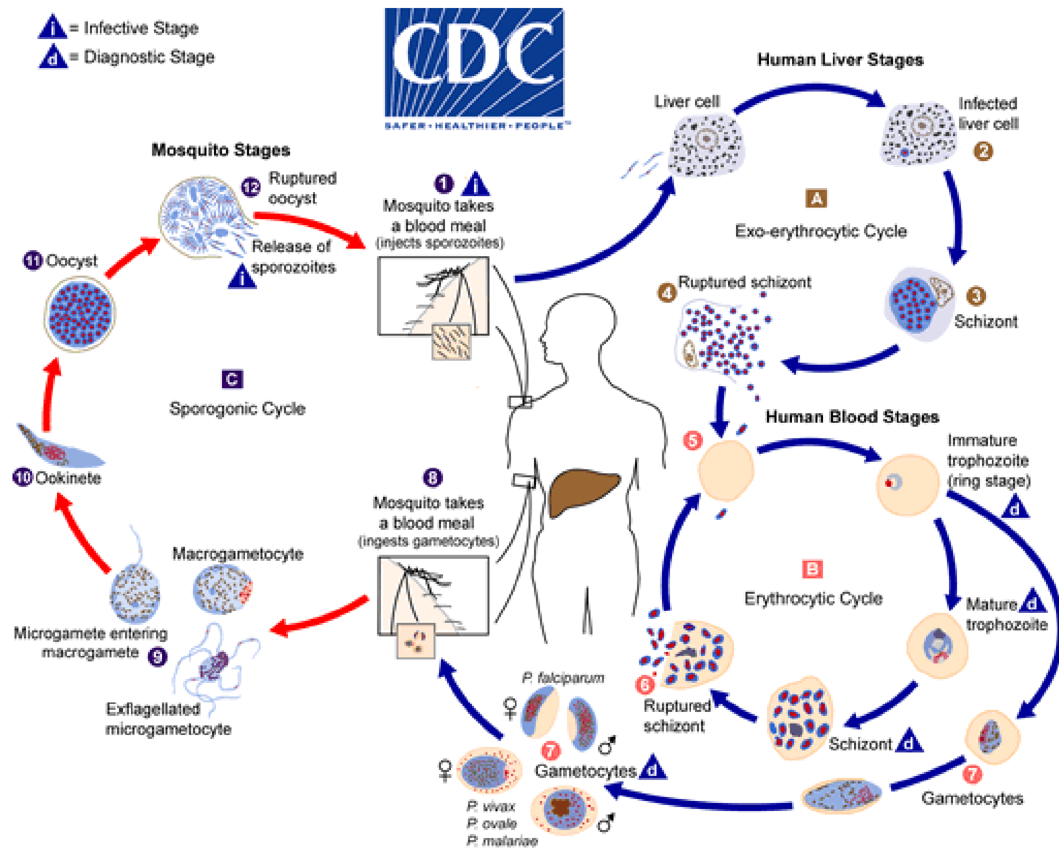
The decrease in malaria incidence and mortality has been credited to advancements in diagnostic techniques, indoor residual spraying, the adoption of treated and long-lasting mosquito nets, and the implementation of artemisinin-based combination therapy (Nkumama et al., 2017). Hence, the decline cannot be attributed to merely a single intervention. Interventions can be effective but are faced with challenges such as drug resistance to asymptomatic infections, which are obscure to treatment and ultimately contribute to the cycle of malaria transmission as well as temporal and spatial variability of the *Plasmodium* species (WHO, 2015). Therefore, increasing access to these interventions and developing tools to address residual transmission and vector control will be the cornerstone of the effort to reduce and eliminate malaria.

Research has shown that pregnant women and children below the age of five years are the most vulnerable groups in areas with high malaria transmission rates (Rogerson et al., 2018). This is because immunity against malaria is typically developed through repeated exposure, yet many children lack this immunity due to limited exposure (Griffin et al., 2015). Conversely, in pregnant women, the accumulation of infected red blood (iRBCs) cells in the placenta leads to recurrent infections when these cells re-enter the bloodstream, and if anti-malarial drugs fail to reach therapeutic levels (Rogerson et al., 2018).

2.2 Malaria Parasites

Approximately 500 *Plasmodium* species infect a range of animals' hosts including mammals, birds, and reptiles. Initially, five species of *Plasmodium* were recognized as causing human malaria: *P. falciparum*, *P. vivax*, *P. malariae, P. ovale*, and *P. knowlesi* (Jain et al., 2014). Recent research has shown that *P. cynomolgi*, typically a parasite found in simians, can equally cause symptomatic infections in humans, thus expanding the list of species responsible for human malaria (Kojom Foko et al., 2023). *P. falciparum* and *P. vivax* are the primary contributors to clinical cases, with *P. falciparum* dominating the sub-Saharan Africa while in regions like Asia and Latin America, *P. vivax* is the main malaria-causing agent (Howes et al., 2016). While historically considered less severe than *P. falciparum* malaria, reports indicate an increase in deaths attributed to *P. vivax* malaria (Douglas et al., 2014).

All *Plasmodium* parasites' life cycle has three main stages: the exo-erythrocytic, erythrocytic, and sporogonic stages, as illustrated in Figure 2.1. Sporogonic stages occur in the vector and the other two in the host (Cox, 2010). The exo-erythrocytic phase begins when a mosquito feeds on a human blood and introduces sporozoites into hepatocytes. Within these cells, the sporozoites multiply asexually and transform into schizonts, each containing numerous merozoites (Soulard et al., 2015). These merozoites enter the bloodstream, and initiate an erythrocytic stage by invading the red blood cells, replicating asexually until they are enough to trigger an immunological response leading to the onset of signs and symptoms of malaria (Gazzinelli et al., 2014). Some merozoites in the blood mature into gametocytes, which are passed to mosquitoes when they bite an infected person, starting the sporogonic cycle in the mosquito's midgut. Conversely, some merozoites mature into schizonts (Soulard et al., 2015).



Source: Centre for Disease Control - Malaria (2020)

Figure 2.1: Illustration of the full life cycle of *Plasmodium* parasites.

2.3 Genotyping of *Plasmodium* Parasites

Molecular characterization of malaria parasites is important in surveillance, understanding the pathogenesis and evolution of drug resistance, identifying novel biomarkers, and studying parasite transmission dynamics and epidemiology (Tadesse et al., 2017). Characterization of *Plasmodium*-specific genes aids in determining the diversity of the *Plasmodium* populations occurring in each geographical location (Soe et al., 2017). Sequence-based biomarkers have been used in genotyping *Plasmodium* parasites for diversity studies as well as drug resistance; for instance, the K13 propeller gene, *pfmdr1,* and *pfcrt* genes are used as molecular markers for artemisinin-resistance in *P. falciparum* species (Zhang et al., 2018).

The advent of whole-genome sequencing (WGS) has resulted in the sequencing the whole *P. falciparum* genome, revealing that *P. falciparum* comprises 14 chromosomes

with a 23-megabase pair genome (Gardner et al., 2002). Genome-wide association (GWAS) mapping has been used to detect important genetic polymorphisms and to identify rare genes associated with malaria infection. For example, GWAS studies were employed to pinpoint uncommon genetic variations linked to severe malaria (Park et al., 2015). Thus, whole-genome sequencing provides an opportunity to study the genetic basis of these parasites and identify the genes involved in disease development and drug resistance (Cowell & Winzeler, 2018).

2.4 Selected Malaria Outcomes

Malaria infections result in various clinical disease outcomes ranging from asymptomatic malaria to severe infection characterized by severe anaemia, kidney failure, metabolic acidosis, hypoglycemia, and cerebral malaria (Boushab et al., 2020). These clinical manifestations depend on previous exposure to malaria, the extent of acquired immunity, the parasite genetic factors, and the host genetic factors (Miller et al., 2002)

2.4.1 Uncomplicated Malaria

Symptomatic malaria can present as severe or uncomplicated malaria. Severe malaria is defined as malaria characterized by vital organ disturbance (WHO, 2000). In contrast, individuals presenting with malaria symptoms without signs of vital organ disturbance are said to have uncomplicated malaria (Taylor-Robinson et al., 2008). Uncomplicated malaria is characterized by symptoms like intermittent episodes of chills and sweats, headaches, fever, anemia, vomiting, and jaundice (WHO, 2000). However, these signs and symptoms are non-specific and may occur in other febrile conditions, both early and late in the infection (Grobusch & Kremsner, 2005). Hence, prompt treatment is necessary to prevent the progression to severe clinical malaria and disease transmission.

WHO recommends artemisinin combination therapy (ACTs) as the primary treatment for uncomplicated malaria (Sinclair et al., 2009). These combinations comprise artemether-lumefantrine, artesunate-amodiaquine, dihydroartemisinin-piperaquine, artesunate-mefloquine, and artesunate-pyronaridine, as outlined by the WHO in 2015.

These combinations generally consist of a potent and fast-acting artemisinin derivative paired with a companion drug that has a slower mode of action.

2.4.2 Asymptomatic Malaria

Asymptomatic malaria refers to the absence of recent symptoms and signs of malaria in the presence of parasitemia upon laboratory confirmation, making it difficult to diagnose asymptomatic infections. No standardized criteria define asymptomatic infection (Kimenyi et al., 2019). Commonly used diagnostic criteria include the absence of malaria signs and symptoms alongside the detection of parasites in dense blood smears and a body temperature below 37.5 degrees Celsius (Laishram et al., 2012). Asymptomatic malaria is widespread not only in regions with high malaria prevalence but also in areas with low transmission rates. Asymptomatic individuals in local settings act as reservoirs for *Plasmodium*, resulting in malaria transmission when they remain untreated. This poses a serious concern as it hinders efforts to combat and eliminate malaria.

Several factors contribute to developing asymptomatic malaria, including naturally acquired immunity and host genetic factors such as hemoglobinopathies (Acquah et al., 2020; Lamptey et al., 2023). Individuals living in regions where malaria is prevalent are often exposed to various parasite strains, leading to the rapid acquisition of immunity to malaria (Magesa et al., 2002). This repeated exposure can enhance the production of anti-inflammatory cytokines and the proliferation of CD4+ T cells, resulting in asymptomatic infections. Gene expression patterns have revealed activation of pro-inflammatory cytokines and tumor necrosis factor-alpha (TNF-α) (Butler et al., 2013). Anti-inflammatory cytokines, however, regulate these responses, dampening parasite clearance and promoting asymptomatic infections (Portugal et al., 2014).

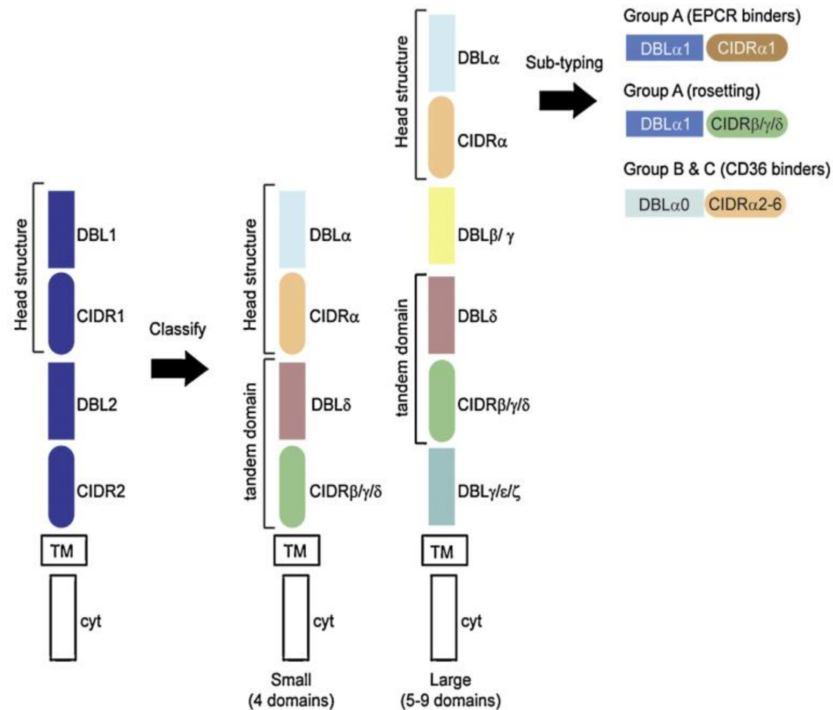
Moreover, chromatin modelling may significantly contribute to the establishment of asymptomatic infections. A study comparing transcriptome profiles in children with different malaria severities in Gabon found that asymptomatic cases exhibited upregulation of RNA processing and nucleotide binding transcripts, indicating active gene regulation through chromatin remodeling (Boldt et al., 2019). Additionally, the study

noted that during asymptomatic infections, the body suppresses immunoglobulin chain transcripts, which contrasts with their activation in cases of uncomplicated infections. This finding aligns with the understanding that chromatin remodeling can lead to the decreased expression of immunoglobulin genes (Bowen & Corcoran, 2008). Chromatin remodelling may be a crucial mechanism for maintaining asymptomatic infections. In summary, the development of asymptomatic malaria involves a complex interplay of genetic, molecular, and immune response mechanisms.

2.5 *Plasmodium falciparum* Virulent Genes and Surface Antigens

About 60 *P. falciparum* virulent *(var*) genes encode variant types of erythrocyte membrane protein 1 (*Pf*EMP1) antigens (Claessens et al., 2014). Roughly 60% of these genes are situated near the *stevor* and *rif* gene families in the sub-telomeric regions of the chromosomes (Kraemer & Smith, 2003). They are classified based on the upstream sequences of the genes as *UpsA, UpsB, and UpsC*, as well as two intermediate groups, A/C & B/C. Upon expression in the host, these *var* genes lead to various clinical outcomes ranging from symptomatic to mild and severe malaria (Bartoloni & Zammarchi, 2012).

The structure of *Pf*EMP1 comprises various partially conserved domains, as depicted in Figure 2.2, such as the Duffy binding-like domains (DBL), a cysteine-rich interdomain region (CIDR), and an N-terminal segment (NTS) (Smith, 2014). Certain DBL domains associated with specific CIDR domains might contribute to severe malaria by disrupting the interaction between activated protein C and the endothelial protein C receptor, leading to the sequestration of infected red blood cells and resulting in severe malaria (Bull et al., 2008).



Source: Smith (2014)

Figure 2.2: The structure of *Pf*EMP1 proteins and the organization of the DBL domains.

2.5.1 *Plasmodium falciparum var* Gene Expression

An association has been noted between *var* gene expression and malaria clinical outcome, where the *Pf*EMP1 type expression determines adherence phenotype and correlates with malaria clinical course (Bertin et al., 2013). Research in Brazil and Mali found that severe non-cerebral malaria was linked to the expression of DBL1α without one to two cysteine residues (Kyriacou et al., 2006). In another example, patients with complicated malaria have shown increased expression of the *var* A transcript and the expression of *var* D (Ariey et al., 2001; Rottmann et al., 2006). Additionally, high parasitemia, coupled with DC6 and DC8 transcripts, correlates with malaria severity in hospitalized cases (Bernabeu et al., 2016).

2.5.2 Single Nucleotide Polymorphisms Within *P. falciparum var* Genes

Single nucleotide polymorphisms (SNPs) are the genetic variations that occur within and outside the coding region and are key to pathogen evolution in response to the

host's stressful conditions, which can influence promoter activity, affecting gene expression, alter mRNA and protein conformations, and subcellular localization of the mRNAs and proteins (Robert & Pelletier, 2018; Shatoff & Bundschuh, 2020). SNPs can be located within virulence genes or genes implicated in immune response, ultimately enhancing virulence or directly implicating the phenotype of susceptibility to infections (Vallejos-Vidal et al., 2020). For instance, in *Bacillus thuringiensis* a specific genetic variation in the ABCC2 loop1 leads to varying levels of toxicity for the CRY1AC toxin (Liu et al., 2018).

In *Plasmodium*, genetic variations within the homology blocks of the *var* genes that could be linked to severe malaria have been identified (Zinder et al., 2017). These variations influence gene expression patterns, resulting in an organism adapting to their immediate environment. Moreover, single nucleotide polymorphisms associated with artemisinin resistance have been identified and validated as drug resistance markers, including *Pf*k13, *Pf*MDR1, and *Pf*crt (Zhang et al., 2018).

2.6 Malaria Transcriptomics

Transcriptomics is widely used in malaria research for gene expression studies. For instance, transcriptomics has been used to reveal that febrile malaria-experienced individuals exhibited a similar transcriptome profile to asymptomatic individuals as compared to naïve individuals (Tran et al., 2016). Transcriptomics has also been used to uncover differences in transcriptomic signatures for cerebral malaria, uncomplicated malaria, and asymptomatic infections (Cabantous et al., 2017). Additionally, transcriptomics has been used to profile human response to malaria, identify molecular markers, understand pathogenicity, and explore the mechanisms behind acquired immunity (Hodgson et al., 2019).

2.6.1 Epigenetic Regulation in Malaria

Epigenetic regulation is defined as heritable phenotypic changes in genes without alteration of the primary DNA sequence of an organism (Holliday, 1987; Cortés & Deitsch, 2017). Epigenetic mechanisms control crucial biological functions of an

organism like embryogenesis and cell differentiation, affirming that epigenetic programming is crucial in transcriptomic heterogeneity (Hanahan, 2022). In *P. falciparum* epigenetics regulates *var* switching (Merrick et al., 2012; Massah et al., 2015; Deitsch & Dzikowski, 2017). This focus stems from their significant role in antigenic variation and virulence. These genes are expressed exclusively, with only one being active at any given time. Histone modifications and noncoding RNA regulation govern this mutual exclusivity (Amit-Avraham et al., 2015). Studies are now expanding our understanding of epigenetics in various other biological functions within malaria parasites. These functions encompass gametogenesis, transport of substances, parasite invasion, and regulation of clonal variants gene expression (Cortés & Deitsch, 2017; Witmer et al., 2020). This underscores the significant impact of epigenetics, which goes beyond mere antigenic variation.

2.6.2 Non-coding RNAs

A significant portion of the genome of an organism undergoes transcription into noncoding RNAs (Kimura, 2020). Non-coding RNAs can be categorized into distinct classes, including PIWI-interacting RNAs, small nuclear RNAs, microRNA, small nucleolar RNAs, long non-coding (lncRNA), and circular RNAs (Li et al., 2021). Circular RNAs and lncRNAs play a crucial role in regulating gene expression through diverse mechanisms. One such mechanism involves serving as scaffolds for protein-protein interactions, influencing the binding of transcription factors, and consequently, regulating specific gene expression patterns (Zhao et al., 2014).

The processes governing gene regulation in *P. falciparum* are not fully understood, but there has been a notable increase in research aimed at uncovering the impact of noncoding RNAs on the parasite's gene regulation (Lodde et al., 2022; Simantov et al., 2022). *Plasmodium* expresses various types of noncoding RNAs, including circRNAs, lncRNAs, and several non-structural RNAs. Although microRNAs have not been found in *Plasmodium*, the parasite employs host microRNAs to modulate its gene expression (Dandewad et al., 2019). Although microRNAs are not found in *P. falciparum*, it is becoming more apparent that long noncoding RNAs have a substantial impact on regulating gene expression.

Further investigation of lncRNAs' regulatory capacity may provide insight into post-transcriptional control and chromatin regulation *in P. falciparum*. Research has demonstrated lncRNAs' regulatory role in the invasion and antigenic variation (Broadbent et al., 2015). Moreover, lncRNAs have been implicated in the *var* genes' mono-allelic expression. The precise regulation mechanisms of this multifamily gene are poorly understood; however, chromatin remodelling, lncRNAs, and histone modifications have been linked with both the activation and suppression of these genes. For example, lncRNA var-AS regulates *P. falciparum var* genes via chromatin modification (Epp et al., 2009).

2.7 Host-parasite Interactions

Host-parasite interactions are changes occurring in either the parasite or the host at the molecular, cellular, or behavioural levels due to the influence of one organism on the other (Smith & Styczynski, 2018). These changes encompass secondary effects as well as more complex outcomes and are crucial for understanding the pathogenesis and *P. falciparum* 's adaptation within the host environment. The molecular-level dynamics of the interaction between hosts and *P. falciparum* parasites have remained elusive. However, modern next-generation sequencing techniques offer opportunities to investigate and clarify the genetic factors that influence these interactions (Greenwood et al., 2016).

Omics have gained significant attention in studying pathogen-host interactions allowing for ontological analysis, phenotypic association studies, and computational modelling to elucidate complex cellular networks (Aderem et al., 2011). Through ontological analysis, gene functions and associated biological processes can be described and important molecular interactions elucidated (Ashburner et al., 2000; Consortium, 2019). Phenotypic association investigates the relationship between biomolecule abundance and the impact of a given trait (Harrison et al., 2012). These computational approaches offer insights into genetic factors driving molecular interactions, and enhance our comprehension of interactions between hosts and *P. falciparum* providing new strategies for malaria control and eradication

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study Site

The samples were collected from participants residing in Kisumu and Kombewa sub counties of Kisumu County at two county referral hospitals namely Kombewa (KOM) and Kisumu (KDH). Kisumu county is situated at the Victoria Lake basin one of the ecological zones of malaria in Kenya. These sites were chosen due to their close proximity to the Malaria Drug Resistance Laboratory in Kisumu City, allowing for easy transportation of samples to the lab within 30 min after collection.

3.2 Ethical consideration

This study was part of ongoing studies, and the use of the samples in this study was approved by the Walter Reed Army Research Institute of Human Use Research Committee, Silver Spring, MD, consent number: WRAIR 2454, and the Kenya Medical Research Institute, Scientific and Ethics Review Unit (SERU) consent number 3628. In addition, informed consent was sought from all participants prior to enrolment into the study, and all data obtained in this study were handled in confidence and laboratory coding was used in the identification of the samples.

3.3 Study Design and Population

This study used a case-control design, recruiting participants who exhibited symptoms of malaria as well as asymptomatic carriers of the disease (Schlesselman, 1982).

3.3.1 Uncomplicated Malaria Participants

Persons from six months of age or older were recruited. The symptomatic malaria samples were collected from six months or older persons who sought treatment for malaria and presented a positive *P. falciparum* on the rapid diagnostic test (RDT) consistent with *P. falciparum* (both test bands +) and microscopy. The RDTs used were compliant with the standard protocol for detecting *P. falciparum* malaria

infections in Kenya. In addition, a cohort of four individuals aged six months or more, who exhibited no malaria symptoms, was included for sample quality assurance.

3.3.1.1 Inclusion Criteria

The participants in the study should have fulfilled the following criteria:

i. Persons aged six months or older attended at a participating MoH hospital facility

ii. Persons who gave written informed consent or assent. For dependent persons, i.e., between six months and 18 years, the guardians had to give consent.

iii. Anyone aged six months or older who had taken antimalarial drugs in the past 14 days from the date of the test but still tested positive for *P. falciparum* (both test bands showing positive) or another species (only the Pan band showing positive) on the *Pf* and/or Pan RDT test.

iv. Pregnant women who gave consent.

3.3.1.2 Exclusion Criteria

Participants were excluded from the study under the following circumstances:

i. Refusal or unwillingness to participate or provide blood samples.

ii. Adults lacking the capacity to give informed consent.

iii. Prisoners, including minors in the Kenya Government Children's Correction and Rehabilitation Centers.

iv. Children under 18 years of age without a parent or legal guardian present.

v. Individuals who had previously participated in the study within the same calendar year surveillance period.

vi. Infants over six months old but less than five kilograms.

vii. Individuals showing severe adverse effects, such as clinical evidence of severe anemia, as determined by the attending medical provider, following the collection of 2.5 ml of blood.

3.3.2 Asymptomatic Participants

Individuals showing no observable signs or symptoms of malaria were recruited for the study. Prior to their participation, a medical history and clinical examination were conducted.

3.3.2.1 Inclusion Criteria

Individuals included in this category had to have met the following criteria:

i. Ability to take oral anti-malarial medications (Coartem and low-dose primaquine)
ii. Persons who gave written informed consent or assent. For dependent persons, i.e., below 18 years, the guardians had to give consent.
iii. For females, they must have been either not capable of bearing children or used suitable contraception methods for 30 days following the administration of Coartem and primaquine.
iv. For males, they must have been willing to ensure that they do not impregnate their partner(s) for a minimum of three months post primaquine treatment.
v. Tested positive for *P. falciparum* by PCR with *ct* value below 31

3.3.2.2 Participant Exclusion Criteria

Participants meeting any of the following criteria were excluded from this study category:

i. If displaying any malaria signs or symptoms
ii. Having contraindications for Coartem and primaquine administration as specified in the respective drug inserts
iii. If pregnant
iv. Use of antimalarial treatment drugs within two weeks preceding the study period
v. Participation in a malaria vaccine study during the study period

vi. Having any other findings that the investigator believed would increase the risk of adverse outcomes from study participation.

3.4 Sampling Criteria and Sample Size

A probability sampling approach was employed to select 50 field samples from two distinct groups: patients exhibiting uncomplicated clinical malaria (n=20) and asymptomatic carriers (n=30). Since previous studies have reported statistically significant differences between groups using a smaller sample size, it was expected that our sample size of 50 would give an 80% statistical power to detect the expected differences. To enhance statistical power, an additional malaria transcriptome dataset was acquired from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) This dataset encompassed categories including data generated from asymptomatic, symptomatic, and naïve symptomatic individuals.

3.5 Retrieval of Malaria Transcriptome Data from GEO

The criteria outlined were used to search for malaria transcriptome datasets.

i. The dataset must have been generated from blood samples obtained directly from the field and not from culture-adapted parasites. This was to ensure that the data reflected the behaviour of *P. falciparum* in its natural environment.

ii. The dataset must have been RNA sequencing data and not from microarray because RNAseq can detect novel, unannotated genes compared to standard microarray technology.

iii. The dataset should have been isolated from asymptomatic malaria carriers, symptomatic or naïve malaria individuals.

3.6 Sample Collection and Storage

Blood samples were collected from consenting individuals visiting Kisumu and Kombewa county referral hospitals between August and December 2021. About 2-3 ml of venous blood sample was collected in acid citrate dextrose (ACD) tubes (Becton-

Dickinson, Franklin Lakes, NJ, USA) using 21-gauge butterfly needles (Becton Dickinson). The blood samples were transported to the Malaria Drug Resistance Laboratory (MDR) at Kisumu at 4– 8°C. Upon arrival at the lab, samples were depleted of the white blood cells using CF11 cellulose. Briefly, the CF11 cellulose was packed dry into 5 ml columns and then wetted with 5 ml Phosphate-buffered saline (PBS). Whole blood underwent centrifugation at 2500 rpm for 5 min, leading to the separation of the serum-containing supernatant. After discarding the supernatant, the pellet containing erythrocytes was resuspended in 200 µl of PBS buffer and passed through the column. The column was subsequently rinsed with PBS until the erythrocytes were visibly eluted. The red blood cells were then centrifuged again at 2500 rpm, for 5 min and the resulting pellet was resuspended in 200 µl of PBS and mixed with 2 volumes of DNA/RNA shield and immediately frozen at -80°C.

3.7 Sample Processing and Sequencing

3.7.1 RNA Extraction from Blood Samples

RNA extraction was carried out using the QIAamp® RNA mini kit from Qiagen as per the manufacturer's protocol with slight modifications. Briefly, the blood samples in RNA shield were thawed to room temperature and vortexed thoroughly until fully homogenized. The homogenized sample was aliquoted into a 2ml Eppendorf tube, and 2% of 20mg/ml proteinase K was added, mixed thoroughly, and incubated a room temperature overnight. The lysate underwent centrifugation at 8700 rpm for 2 min then subsequently transferred to a new Eppendorf tube. Subsequent steps were carried out following the QIAamp kit manufacturer's instructions.

3.7.2 Ethanol Precipitation of RNA

To achieve the desired RNA purity and concentration, the extracted RNA was precipitated overnight in 0.1 volumes of 3M Sodium acetate and three volumes of absolute ethanol. Subsequently, the precipitated RNA underwent washing with 75% ice-cold ethanol and was then reconstituted in 35 microliters of 1X T.E buffer.

3.7.3 RNA Quantification and Integrity Analysis

The RNA quantity was quantified using the Qubit 3.0 fluorometer and the Qubit RNA HS Assay (Thermo Fisher Scientific). The quality of the purified RNA was evaluated using the TapeStation 4200 system with RNA HS high-sensitivity screen tapes. RNA samples with an RNA integrity number (RIN) exceeding 4.0 were considered for ribosomal depletion.

3.7.4 Molecular Detection of *Plasmodium* Species

The presence of *Plasmodium* parasites in each sample was confirmed using the qualitative real time PCR; hence, primers targeting specific regions of the 18s rRNA were used to amplify the region. All samples with ct values of 32 and below were considered for further analysis. The assay was carried out on a real-time PCR with RNAS P as an endogenous control. The samples were assayed in a total reaction volume containing 7.5 µl 2x PCR buffer, 0.4µl of the 10 µM *Plasmodium* forward and reverse primers, 0.6µl of the reverse transcriptase enzyme, 0.4 µl RNAS P forward & reverse primer and 1.7 µl of nuclease-free water (Table 3.1). TaqMan probes labelled with the 5′ reporter dyes FAM and VIC were used for *Plasmodium* and RNAS P detection respectively (Table 1). The PCR was performed with the following cycling parameters: initial denaturation at 95°C for 5 min, followed by 40 cycles of denaturation at 95°C for 15 seconds, and an annealing/extension step at 60°C for 1 min.

Table 3.1 Primer and Probe Sequences for the qualitative real PCR Reaction

Probe/Primer	Reporter	Sequence	Quencher
RNASEP Probe	VIC	TGCGCGGACTTGTGGA	MGB
RNASEP FWD	N/A	TGTTTGCAGATTTGGACCTGC	N/A
RNASEP RVS	N/A	AATAGCCAAGGTGGAGCGGCT	N/A
PLU Probe	FAM	ATGGCCGTTTTTAGTTCGTG	MGB
PLU FWD	N/A	GCTCTTTCTTGATTTCTTGGATG	N/A
PLU RVS	N/A	AGCAGGTTAAGATCTCGTTCG	N/A

3.7.5 *Plasmodium* Speciation Assay

This assay utilized consensus primers designed specifically for four targets, namely *P. malariae*, *P. falciparum*, *P. ovale curtisi*, and *P. ovale walkeri* (Table 3.2). Given the uniqueness of the probes, the assay was multiplexed. Individual amplification was conducted in a 12 µl final volume, comprising 7.5 µl of 10x RNA buffer, 400nM Forward primer, 400nM Reverse primer, 400nM probe primer, 0.5µl nuclease-free water, and 0.6µl of reverse transcriptase enzyme and 2.2µl of the nucleic material in a 96, 0.1 ml real-time PCR plate. The thermal cycling protocol was as follows: Pre-PCR incubation at 50°C, initial denaturation at 95°C for 5 min, followed by 40 cycles of denaturation at 95°C for 15 seconds and annealing/extension at 60°C for 60 seconds. Individual reactions containing the three species were included as positive controls, and a single reaction without nucleic acid was included as a negative control. Samples positive for *P. falciparum* by confirmatory PCR and with a threshold below 34 ct were considered for further analysis.

Table 3.2. Consensus primers for Plasmodium speciation assay

Probe/Primer	Reporter	Sequence	Quencher
FAL Probe	FAM	CATAACAGACGGGTAGTCAT	BHQ-1
FAL FWD	N/A	ATTGCTTTTGAGAGGTTTTGTTA CTTT	N/A
FAL RVS	N/A	GCTGTAGTATTCAAACACAATGA ACTCAA	N/A
OVA-c probe	FAM	CCTTTTCCCTATTCTACTTAATTC GCAATTCATG	BHQ-1
OVA-c FWD	N/A	TTTTGAAGAATACATTAGGATAC AATTAATG	N/A
OVA-c RVS	N/A	CATCGTTCCTCTAAGAAGCTTTA CAAT	N/A
Ovav probe	VIC	CCTTTTCCCTTTTCTACTTAATTC GCTATTCATG	BHQ-1
OVA-Fv,	N/A	TTTTGAAGAATATATTAGGATACA TTATAG	N/A
OVA-Rv	N/A	CATCGTTCCTCTAAGAAGCTTTA CAAT	N/A
MAL3P	VIC	TGTTCAAAGCAAACAGTTAAAA CA	BHQ-1
MAL3F	N/A	GCATGGGAATTTTGTTACTTTGA	N/A
MAL5R	N/A	ATGCTGTAGTATTCAAACACAGA AAC	N/A

3.7.6 Genomic DNA Digestion and mRNA Enrichment

To enrich the mRNA, the Ribo-Zero Gold epidemiology rRNA depletion kit was used following the guidelines provided by the manufacturer. The process involved performing ribosomal RNA depletion in a 96-deep well plate. Each reaction comprised 26µl of total RNA, 10µl of Ribo-zero removal solution, and 4µl of Ribo zero reaction buffer, which were mixed and allowed to incubate at room temperature for 5 min. The treated RNA was then transferred to a fresh deep well plate containing washed magnetic beads. After thorough mixing and another 5-minute incubation, the reaction

was vortexed at a medium speed of 2500 rpm for 10 seconds followed by a 5 min incubation at 50°C. Subsequently, the reaction was promptly placed on a magnetic stand until the solution was clear.

A total of 85µl of the supernatant was transferred into a new labelled deep well plate and purified using the AMPURE RNA XP purification beads following the manufacturer's instructions. To assess potential RNA loss during the process, the purified RNA was quantified on Qubit 3.0 fluorometer. To eliminate any genomic DNA contamination, the ribosomal depleted RNA was subsequently treated with DNAse I® (Invitrogen) and purified again using Ampure RNA purification beads as per the manufacturer's guidelines.

3.7.7 Library Preparation

Library preparation was done using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina. Guided by the manufacturer's instructions, RNA was fragmented, and double-stranded cDNA fragments were generated via reverse transcription. The mRNA fragmentation was empirically optimized at 94°C for 10 min to generate 200 bp fragments. Adapter ligated DNA was size-selected using 0.5 times sample volume AMPure XP beads in a buffer to bind and remove large fragments. Then 0.8 times sample volume beads buffer (i.e., 0.3 times sample volume fresh beads) was added to bind and retain intermediate DNA fragments. Libraries of approximately 330 bp that contained approximately 200 bp inserts were generated by this process.

Ligated RNA was amplified in a 50 µl reaction containing NEB User enzyme, 3µl (for excision of uracil in loop adapters), the NEBNext Universal PCR primer for Illumina, and NEBNext Indexed primers for Illumina both at 0.5 µM, 0.3 mM each dNTP, 1 x Kapa HiFi buffer (containing TMAC) (Kapa Biosystems, Roche USA) and 1 unit of Kapa HiFi. The PCR reaction was incubated at 37°C for 15 min, followed by denaturation at 98°C for 1 min. This was followed by 15 cycles of denaturation at 98°C for 10 seconds, annealing at 65°C for 1 min, and extension at 65°C for 5 minutes. Following the manufacturer's instructions, the RNASeq libraries were purified using the Agencourt® AMPure beads. Libraries were quantitated by the Qubit using the

dsDNA HS Assay kit. The libraries' quality was assessed on a TapeStation 4200 system using a D1000 High Sensitivity DNA tapescreen. The libraries were pooled in equimolar quantities before RNA sequencing on the Illumina® MiSeq™.

3.8 Transcription Patterns of *P. falciparum* Parasites During Asymptomatic Malaria Infections

3.8.1 Quality Assessment of the Raw Sequence Data

The raw RNA reads were assessed for poor quality reads and adapter contamination using FastQC v 0.11.9. Following this, Trimmomatic v0.39 was employed to remove adapter sequences that were attached to the reads (Bolger et al., 2014). Any reads with an ambiguous sequence content ("N") exceeding 5% were excluded from further analysis. Using a *P. falciparum* 3D7 reference genome (GCF_000002765.5), the processed reads were then aligned using HISAT2 tool (Kim et al., 2019). Subsequently, the resulting SAM files underwent sorting and conversion into binary BAM format using Samtools.

3.8.2 Gene Expression Analysis

Using the processed bam files, data on gene-level counts was obtained using the Rsubread FeatureCounts software version 1.6.1 (Liao et al., 2014). Following the gene-level quantification, the DESeq2 package in R was used to identify genes that showed differential expression using default parameters (Love et al., 2014). Genes were deemed differentially expressed if their adjusted *p*-value (adjusted for False Detection Rate) < 0.05 (Benjamini & Hochberg, 1995), and exhibited a log2 fold-change (log2FC) of at least ± 1 between the compared groups. These genes were then selected for further analysis.

3.8.3 Gene Ontology and Pathway Analysis

The analysis of functional annotation for the differentially expressed genes was conducted using ClusterProfiler 4.6.0, an integral part of the R statistical analysis tool (Yu et al., 2012). The genes underwent gene ontology annotation using the GO

database followed by pathway analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, with a significance threshold set at $p < 0.05$.

3.8.4 Long noncoding RNA Characterization

The aligned reads were assembled and merged using StringTie (version 1.0.1) using specific arguments (-f 0.01 -c 0.01, -m). Transcripts with lengths below 200 nt and fewer than 2 exons were excluded. Following this, gffCompare v0.11.2 was used to annotate the transcripts (Pertea & Pertea, 2020), followed by use of custom scripts to select transcripts classified as "x," "i," "u," "e," and "o," while eliminating undesired transcripts. Transcripts with fewer than 2 exons were also eliminated during this process.

Analysis of the retained transcripts' protein-coding potential utilized CPC, PLEK, and FEEnc software tools (Kong et al., 2007; Li et al., 2014; Wucher et al., 2017). Transcripts lacking protein-coding potential were kept for further analysis. This process was repeated to confirm the protein-coding ability of the remaining transcripts, with those lacking such ability retained for subsequent analysis.

3.8.5 Analysis of Differentially Expressed Long noncoding RNA

The study employed DESeq2 for log2FC computation of transcripts per million pertaining to lncRNAs. DE-lncRNAs were identified as those with a log2FC ≥ 2 with adjusted *p*-value < 0.05. To explore the regulatory role of lncRNAs on immune evasion, an interaction analysis was conducted between the 28 immune evasion genes and the top 15 upregulated lncRNAs using LncTar software.

3.9 The Expression of the *var* Genes Subsets in ASM versus UM Carriers

3.9.1 Extraction of *PfEMP1* Encoding and Regulator Genes

The genomic coordinates of the *P. falciparum* virulent genes as well as their regulator proteins' coding genes were obtained from the NCBI archive and used to prepare a bed

file. The bed file was used to extract all reads from the sorted bam files that aligned to the genomic coordinates using bed tools intersect (Quinlan, 2014).

3.9.2 *Plasmodium falciparum* Virulence Genes Expression Analysis

The obtained reads files were assembled using Cufflinks v2.2.1 and subsequently merged using Cuffmerge v1.0.0. The obtained gtf file was used to assign genomic features to the extracted reads in Featurecounts. Analysis of gene expression patterns was conducted using the DESeq2 package to elucidate transcription patterns between the groups. To investigate the *var* gene regulation, we explored the expression patterns of genes coding for the *var* gene regulator proteins, including KAHRP, KAHSP40, HSP70-x, Mauler's cleft proteins (SBP1), and REX1.

3.10 Single Nucleotide Polymorphisms Within *var* Genes Associated with Asymptomatic Malaria

3.10.1 Alignment to the pf3D7 Reference Genome

The Bwa-mem tool was employed to map the unprocessed reads from the 50 samples against the GCF_000002765.5 reference genome using the parameters -c 100 -T 50 (Li, 2014). Using Bedtools and prior bed file, the genomic coordinates covering the *var* genes were used to extract all reads that mapped to the *var* genes' genomic coordinates as previously described in section 3.9.1. The obtained bam files were sorted and indexed in Samtools. The sorted bamfiles were used for variant calling. BCFTools were used to call variants using default parameters. Given that *P. falciparum* has a haploid genome, the ploidy argument was specified. The raw variants were filtered using specific criteria: maximum missing rate of less than 20%, quality score (QUAL) > 30, Depth (DP) > 100, and a minor allele frequency (MAF) greater than 0.05. SnpEff v5.2 software was used annotate the variants (Cingolani et al., 2012).

CHAPTER FOUR

RESULTS

4.1 Demographic Characteristics of Study Participants

The gender distribution was 50% male and 50% female for asymptomatic cases and 55% male and 45% female for uncomplicated cases. The age distribution for both conditions was comparable. The average age of asymptomatic carriers was 39.03 years with an age range from 25 to 55 years, whereas participants with uncomplicated malaria had an average age of 33.15 years with an age range from 12 to 54 years.

4.2 *Plasmodium* Detection and Speciation of Malaria Samples

All 50 samples were positive for the endogenous RNAS P control. While all 30 asymptomatic samples were positive for *Plasmodium* via PCR detection at a ct threshold of < 35. Eighteen out of the 20 uncomplicated malaria samples were positive, whereas two were undetermined and were excluded from further analysis. These samples were subjected to a speciation assay to establish positive *P. falciparum* samples only. Species information on 3 of the 18 uncomplicated malaria samples were undetermined and were also excluded from further analysis.

In uncomplicated malaria cases, the parasite prevalence for each species was observed as 53.3% (8 out of 15) for *P. falciparum* (*Pf*) and 6.7% (1 out of 15) for *P. malariae (Pm)*. Notably, while the prevalence of pure *P. falciparum* cases remained notably high, there was also a significant presence of mixed species (*P. falciparum* and *P. malaria*) samples, with 6 cases (40% of the total 15 *Plasmodium*-positive samples).

In asymptomatic malaria carriers, the parasite prevalence was 44.4% (24 out of 30) for *P. falciparum* and 6.7% (2 out of 30) for *P. malariae*. Similar to uncomplicated malaria infections, the prevalence of pure *P. falciparum* species was notably high, constituting 80% of all samples. The prevalence of mixed species (*Pf* and *Pm*) in asymptomatic infections was comparatively low at 13.3% (4 out of 30) compared to 40% (6 out of 15) in uncomplicated cases. Samples positive for *P. falciparum* with a ct threshold < 32 were included for further analysis. Combining the eight *P. falciparum*-positive

samples in uncomplicated malaria with the 24 *P. falciparum*-positive samples in asymptomatic malaria, along with the 6 and 4 samples with mixed infections yielded 42 *P. falciparum* positive samples.

Evaluation of the population gametocyte carriage prevalence between the asymptomatic infections and uncomplicated malaria infection revealed that the prevalence was at 33.3% (10 out of 30) for asymptomatic malaria carriers, which was significantly different (p = 0.028) to 72.2% (13 out of 18) for uncomplicated malaria infection. Gender did not show a significant disparity in gametocyte carriage (p = 0.555) (Table 4.1).

Table 4.1. Gametocyte carriage in both asymptomatic and uncomplicated malaria samples.

Variables	Category		Gender	
	Asymptomatic malaria	**Uncomplicated malaria**	**Male**	**Female**
Positive	33.3%	72.2%	52%	47.9%
Negative	66.7%	27.8%	48%	52.1%
$\chi 2$ (p-value)	4.844 (0.028)		0.349 (0.555)	

4.3 RNA Sequencing

4.3.1 RNA Quantification and Integrity Analysis

The extracted RNA from the 42 *P. falciparum* positive samples was quantified and quality evaluated to filter low-quality samples. Across all study samples, the RNA exhibited relatively high quality, with RNA integrity numbers ranging from 5.6 to 8.0. The concentrations ranged from 6.1 to 42.2 ng/µl. A representative of the spectrophotometric and integrity results is shown in Table 4.2. Full

data can be accessed through https://github.com/JGisaina/supplimentary_files

Table 4.2. A representative of the quantity and the quality results of RNA isolated from both asymptomatic and uncomplicated malaria samples.

Sample ID	RNA Concentration	RNA integrity number (RIN)
ASM 181	26.6 ng/ ul	7.1
ASM 183	42.2 ng/ µl	7.3
ASM 184	19.1 ng/ µl	7.1
ASM 189	23.5 ng/ µl	6.4
ASM 227	28.5 ng/ µl	7.6
ASM 265	18.7 ng/ µl	8.3
ASM 298	24.8 ng/ µl	7.5
ASM 309	4.1 ng/ µl	6.9
UM0040	8.98 ng/ µl	5.4
UM0041	19.3 ng/ µl	5.7
UM0045	29.2 ng/ µl	7.2
UM0046	23.9ng/ µl	6.4
UM0047	11.8 ng/µl µl	5.6
UM0048	26.2 ng/ µl	7.3
UM0049	16.1 ng/ µl	8.0

4.3.2 Library Preparation

A total of 42 cDNA libraries were generated representing 14 samples from uncomplicated malaria samples and 28 samples from asymptomatic carriers. The concentration of the cDNA libraries ranged from ~ 10.5 to ~ 52 ng/µl with an average of 16 ng/µl . Additionally, the size of the DNA libraries acquired was assessed using Tape station 4200 systems, revealing an average length of 259 base pairs (bp). The individual fragments varied in length, ranging from 212 bp to 450 bp. A representative of the libraries' characteristics is shown in Table 4.3. Full data can be accessed through https://github.com/JGisaina/supplimentary_files

Table 4.3. Representative samples showing the characteristics of the sequence libraries.

Sample ID	Library concentration	Library fragment size (bp)
ASM 181	18.5 ng/µl	261
ASM 183	24.0 ng/µl	328
ASM 184	19.3 ng/µl	257
ASM 189	20.2 ng/µl	384
ASM 227	18.6 ng/µl	417
ASM 265	25.7 ng/µl	287
ASM 298	10.7 ng/µl	224
ASM 309	15.2 ng/µl	234
UM0040	19.3 ng/µl	246
UM0041	7.76 ng/µl	298
UM0045	34.9 ng/µl	342
UM0046	29.7ng/µl	318
UM0047	31.0 ng/µl	444
UM0048	31.2 ng/µl	439
UM0049	38.9 ng/µl	289

4.4 Characteristics of the RNA Sequence Datasets

4.4.1 Quality Assessment of the Raw Sequence Dataset

The sequencing of the cDNA libraries from the 42 samples on the Miseq sequencer generated a total of 43790140 reads with an average of 486557 reads per forward (R1) and reverse (R2) reads of each sample. Trimmomatic 0.39 software was used to trim poor quality reads and remove adapter sequences. All samples with less than 400,000 reads for both R1 and R2 were excluded from further analysis, resulting in 15 samples (ASM, n=8, and UM, n= 7) with a total of 32165536 reads and a mean of 642009 reads per forward and reverse read (Table 4.4); (Appendix 2).

Table 4.4. Characteristics of the raw sequence dataset

	No. of raw reads (n=42)	No. of filtered reads
Total reads	43,790,140	32,165,536
Min	222,404	419,537
Max	1,830,919	183,0919
Mean	486,557	642,009

4.4.2 Dataset Retrieval from NCBI

Uncomplicated malaria transcriptome dataset was retrieved from the GEO database under the GEO Series accession number PRJEB21707. Similarly, the transcriptome datasets for asymptomatic and naïve malaria were obtained from the GEO database under accession number GSE148125. The latter dataset was derived from *P. falciparum* parasites isolated from venous blood samples obtained from Malian children persistent with subclinical malaria during the dry season (n= 11) and children of the same age with their first clinical malaria during the wet season (n= 12), representing naïve malaria infections while the PRJEB21707 dataset was generated from venous blood samples collected from individuals presenting with uncomplicated malaria in a healthcare facility in Timika, Indonesia (n=12). The dataset was processed as described in section 4.4.1.

4.4.3 Alignment and Transcript Assembly

This process resulted in an average of 8.4 million transcripts per sample. Among these, 295.7 million transcripts (35.91%) were mapped to unique regions of the reference genome, and 5,201,274 (0.5%) mapped as singleton reads (Table 4.5). The proportion of transcripts mapping to the reference genome varied from 1.7% to 91.1%, with asymptomatic samples having the lowest mapping rate, averaging 1.95 million reads per sample and comprising only 7.2% of the total mapped reads

Table 4.5: Alignment Summary Statistics

Variable	**Asymptomatic n = 19**	**Symptomatic UM = 19, NM = 12**	**Total reads n = 50**
Total reads (in a million reads)	446.9	261.4	823.5
Mapped reads (in a million reads)	21.5	203.2	295.7
Itself and mate mapped (in a million)	15.0	170.5	240.4
Singletons (in a million)	0.7	2.9	5.2
Mapped to a different location (in million)	0.7	2.9	1.8

4.5 Transcription Patterns of *P. falciparum* Parasites During Asymptomatic Malaria Infections Compared to Symptomatic Infections

4.5.1 Comparative Analysis of DEGs Between ASM Carriers Versus UM

Thirty-eight samples (ASM n = 19, UM n = 19) were analyzed to uncover the differentially expressed transcripts in asymptomatic versus uncomplicated malaria infection. The principal component analysis revealed the clustering of the samples (Figure 4.1). PC1 and PC2 explained 78.81% and 4.28% of the variability, respectively.

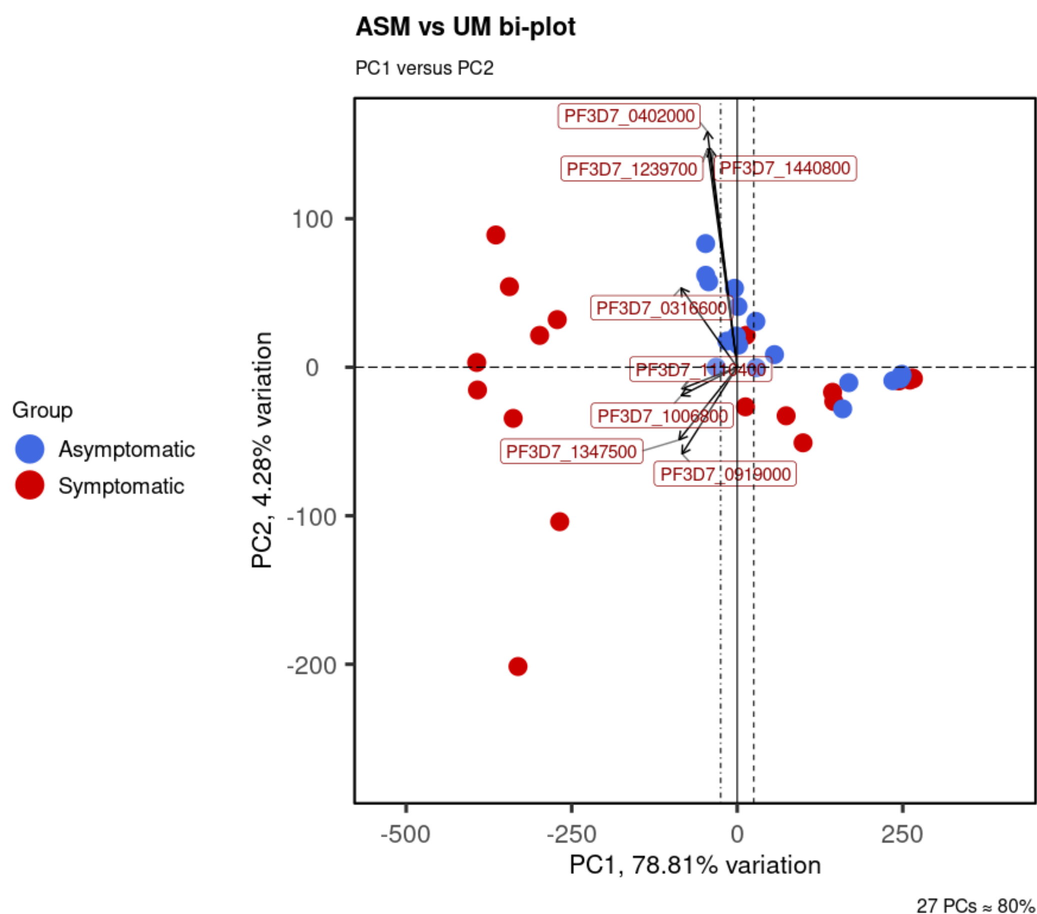


Figure 4.1. Illustration of a principal component analysis demonstrating the variation and clustering observed among samples from individuals with ASM compared to those with UM. The red points correspond to asymptomatic samples, whereas the blue points represent samples from individuals with uncomplicated malaria. Highlighted genes indicate the transcripts contributing significantly to the observed variation.

The analysis of gene expression differences between asymptomatic carriers and individuals with uncomplicated malaria using DESeq2 revealed 3094 genes showing distinct expression patterns. Among these genes, 22 exhibited significant upregulation while 3072 genes were significantly downregulated (adjusted p-value < 0.05 and log2 fold-change of ± 1). These results are visualized in a volcano plot (Figure 4.2).

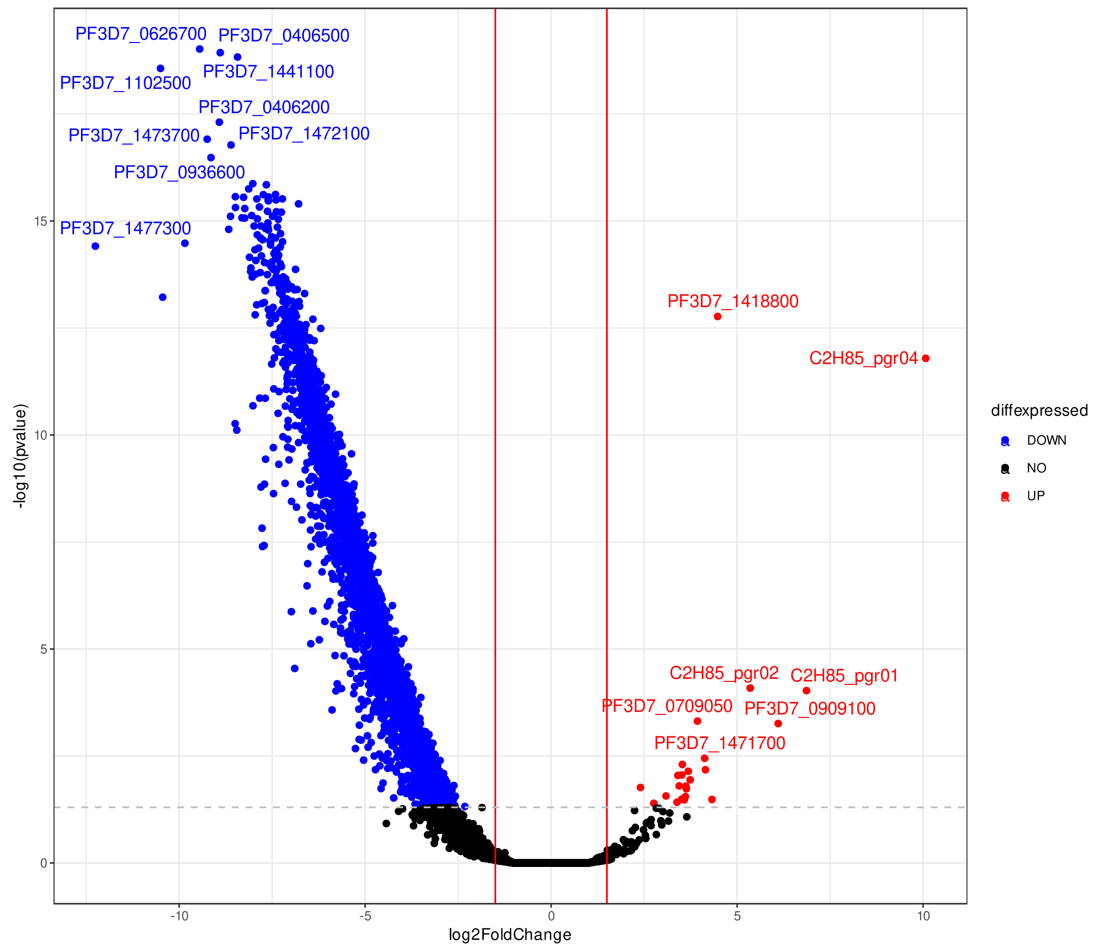


Figure 4.2. A volcano plot showing differentially expressed genes in ASM vs. UM as a function of fold change ± 1 and mean expression in asymptomatic vs uncomplicated malaria. The horizontal axis indicates the log2 mean expression level, while the vertical axis shows the log2 transformed fold change. The red dots represent genes that exhibited a notable increase in expression, while the blue dots indicate genes that showed a decrease in expression. Black dots represent genes that did not show significant differential expression.

Among the top ten upregulated genes were the procollagen lysine 5-dioxygenase, putative, and several uncharacterized genes (Table 4.6). The study further revealed downregulation of genes coding for exported proteins of unknown functions, helical interspersed subtelomeric (*phist*) genes, and transcription factors with APC2 domains (Table 4.7)

Table 4.6. Top 10 upregulated genes in asymptomatic compared to uncomplicated malaria at padj value of <0.05 and log2FoldChange of ± 1.

Gene ID	Description	log2FoldChange	Padj
PF3D7_0909100	Conserved *Plasmodium* membrane protein, unknown function	6.104022798	0.000995
PF3D7_1418800	Signal recognition particle RNA	0.471846971	5.08E-12
PF3D7_0930000	Procollagen lysine 5-dioxygenase, putative	1.16020054	0.01047607
PF3D7_1471700	Conserved *Plasmodium* protein, unknown function	1.073398335	0.005836451
PF3D7_0709050	Small nucleolar RNA	0.841135076	0.000879125
PF3D7_1244900	Conserved *Plasmodium* protein, unknown function	1.081599756	0.017384939
PF3D7_1141900	Inner membrane complex protein 1b, putative	0.999906278	0.01131796
PF3D7_0933300	Conserved *Plasmodium* protein, unknown function	1.11814862	0.027352954
PF3D7_1471600	Conserved *Plasmodium* protein, unknown function	1.088459295	0.024000369
PF3D7_0518800	Secreted ookinete protein, putative	1.188132114	0.040634161

Table 4.7. Downregulated genes in asymptomatic versus uncomplicated malaria infections at padj value of <0.05 and log2FoldChange of ± 1.

Name	Description	log2FoldChange	Padj
PF3D7_1472200	Histone deacetylase, putative	-8.488747026	7.07E-14
PF3D7_1222600	Transcription factor with AP2 domain(s)	-8.494998835	4.62E-10
PF3D7_1472100	Protein transport protein YIP1, putative	-8.60540332	1.05E-14
PF3D7_0113300	*Plasmodium* exported protein (hyp1), unknown function	-8.618782279	1.10E-13
PF3D7_1115500	Transcription factor with AP2 domain(s), putative	-8.666593491	1.64E-13
PF3D7_0406500	Conserved *Plasmodium* protein, unknown function	-8.89502762	2.16E-16
PF3D7_0406200	Sexual stage-specific protein precursor	-8.918906547	4.31E-15
PF3D7_0936600	Gametocyte exported protein 5	-9.143857551	1.81E-14
PF3D7_1473700	Nucleoporin NUP116/NSP116, putative	-9.246070365	8.97E-15
PF3D7_0626700	Conserved protein, unknown function	-9.447396021	2.16E-16
PF3D7_1476600	*Plasmodium* exported protein, unknown function	-9.845130802	2.74E-13
PF3D7_0937200	Lysophospholipase, putative	-10.44216665	2.41E-12
PF3D7_1102500	*Plasmodium* exported protein (PHISTb), unknown function	-10.50037496	2.99E-16

4.5.2 Comparative Analysis of DEGs Between ASM Carriers Versus Naïve Malaria Infections

Thirty-one samples were analyzed to determine the differentially expressed transcripts in asymptomatic and in naïve (NM) malaria infection. The variation between the groups and clustering of samples were evaluated using the principal component analysis. It was found that PC1 and PCA2 accounted to 78.81% and 4.28% of the

variation, respectively (Figure 4.3).

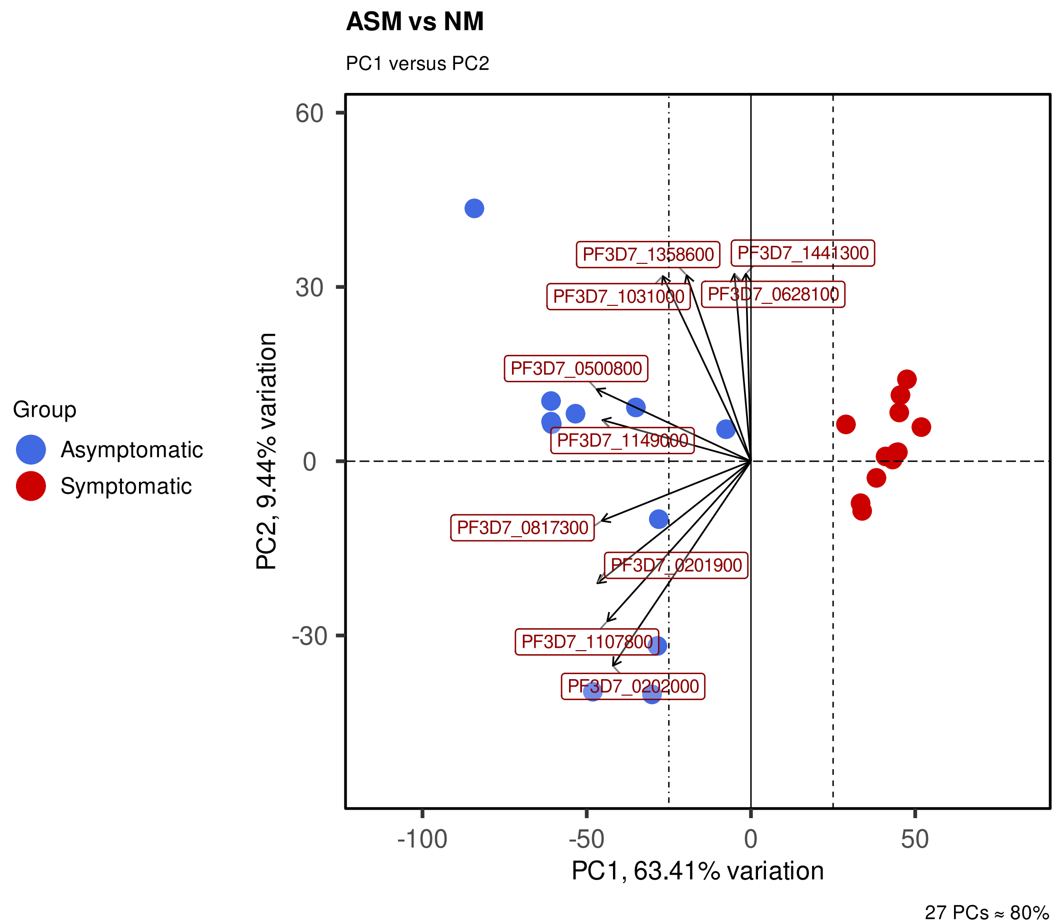


Figure 4.3. Illustration of a principal component analysis demonstrating the variation and clustering observed among samples from asymptomatic individuals versus naïve malaria individuals. The red points correspond to asymptomatic samples, whereas the blue points represent samples from individuals with naive malaria. Highlighted genes indicate the transcripts contributing significantly to the observed variation.

The analysis revealed 1796 genes showing notable variations in expression levels between individuals with asymptomatic malaria and those naïve malaria, using a padj cut-off <0.05 and a log2FoldChange of ± 1. Among these, 1553 genes showed upregulation, while 220 genes showed downregulation (Figure 4.4).

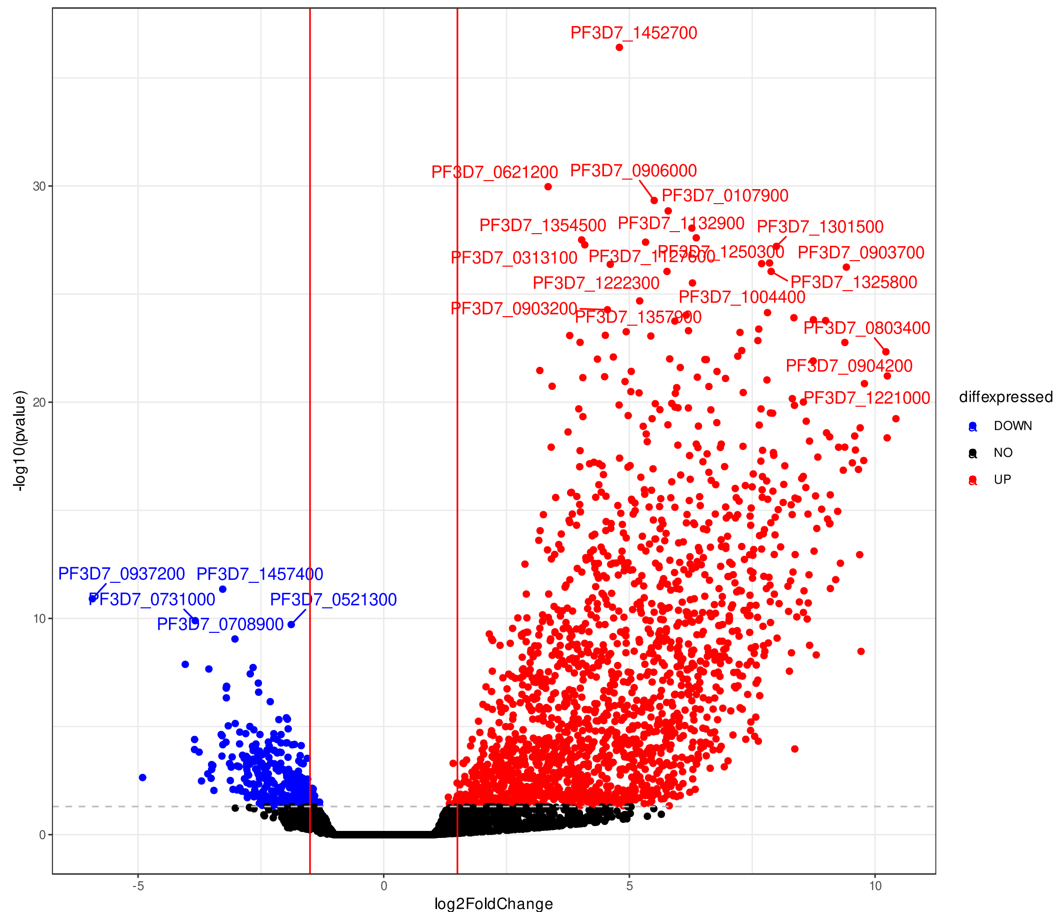


Figure 4.4. A volcano plot illustrating differentially expressed genes in ASM vs NM, comparing fold change ± 1 and mean expression between asymptomatic malaria and naive malaria. The horizontal axis indicates the log2 mean expression level, while the vertical axis shows the log2 transformed fold change. The red dots represent genes that exhibited a notable increase in expression, while the blue dots indicate genes that showed a decrease in expression. Black dots represent genes that did not show significant differential expression.

Among the top 10 upregulated genes were genes coding for conserved proteins of unknown functions, kinases, histone-lysine N-methyltransferase, H3 lysine-4 specific, and DNA repair and recombination elements (Table 4.8).

Table 4.8. Top 10 upregulated genes in asymptomatic versus naïve malaria infections at padj value of <0.05 and log2FoldChange of ± 1.

Gene ID	Description	log2FoldChange	Padj
name	description	10.42403924	3.84E-18
PF3D7_0930000	procollagen lysine 5-dioxygenase, putative	10.25074026	6.77E-20
PF3D7_0904200	conserved *Plasmodium* protein, unknown function	10.24590604	2.55E-17
PF3D7_1471700	conserved *Plasmodium* protein, unknown function	10.2208671	6.85E-21
PF3D7_0803400	DNA repair and recombination protein RAD54, putative	9.783916192	1.34E-19
PF3D7_1221000	histone-lysine N-methyltransferase, H3 lysine-4 specific	9.769334441	2.33E-16
PF3D7_0407800	conserved *Plasmodium* protein, unknown function	9.714681178	3.13E-08
PF3D7_0504500	conserved *Plasmodium* protein, unknown function	9.696912947	9.36E-18
PF3D7_0607700	conserved *Plasmodium* protein, unknown function	9.687006001	2.21E-12
PF3D7_0515400	conserved *Plasmodium* protein, unknown function	9.662484443	5.38E-16

4.5.3 Function Classification by Gene Ontology of Differentially Expressed Genes in both ASM vs. UM and ASM and NM.

To better understand the pathogenesis of asymptomatic malaria, enrichment analysis was performed for the 3094 differentially expressed genes in asymptomatic malaria compared to uncomplicated malaria, and 1753 DEGs in asymptomatic malaria compared to uncomplicated samples. A total of 281 ontology terms were assigned to the 3094 differentially expressed genes in asymptomatic versus uncomplicated malaria, categorized into three main groups: 62 molecular functions, 93 cellular

processes, and 126 biological processes. The list of the enriched GO terms can be accessed from https://github.com/JGisaina/supplimentary_files.

In the molecular functions category, several types of binding activities were predominant, including binding (GO:0005488), GO:0003723 (RNA binding), GO:0003676 (nucleic acid binding), GO:0097159 (organic cyclic compound binding), GO:0005515 (protein binding), and GO:1901363 (heterocyclic compound binding), as detailed in Table 4.9.

Table 4.9. Top 10 Enriched molecular function terms in asymptomatic vs. uncomplicated malaria. (A representative of the enriched terms)

ID	Description	Set size	Enrichment score	*p*-value
GO:0005488	Binding	593	-0.295721608	1.00E-04
GO:0003676	nucleic acid binding	278	-0.321373627	1.00E-04
GO:0097159	organic cyclic compound binding	306	-0.328338972	0.0001
GO:1901363	heterocyclic compound binding	305	-0.326720406	0.0001
GO:0005515	protein binding	304	-0.283752956	0.0001
GO:0003723	RNA binding	234	-0.36150304	0.0001
GO:0003729	mRNA binding	64	-0.35080083	0.000104
GO:0036094	small molecule binding	37	-0.463716415	0.00011
GO:0032553	ribonucleotide binding	35	-0.467502535	0.000111
GO:0017076	purine nucleotide binding	34	-0.47794906	0.000111

In the biological processes category, most of the terms were supressed (Table 4.10). The most predominant terms with a set size greater than 350 genes were: GO:0009987 (cellular process), GO:0008150 (biological process), GO:0044238 (metabolic process), GO:0071704 (organic substance metabolic process), and GO:0043170 (macromolecule).

Table 4.10. Representative of enriched biological process terms in asymptomatic vs. uncomplicated malaria.

ID	Description	Set size	Enrichment score	*p*-value
GO:0043170	macromolecule metabolic process	362	-0.257742879	1.00E-04
GO:0006807	nitrogen compound metabolic process	397	-0.25001259	1.00E-04
GO:0044238	primary metabolic process	408	-0.24556547	1.00E-04
GO:0044237	cellular metabolic process	389	-0.243225013	1.00E-04
GO:0008150	biological process	767	-0.231436644	1.00E-04
GO:0008152	metabolic process	444	-0.237733436	1.00E-04
GO:0071704	organic substance metabolic process	434	-0.234885969	1.00E-04
GO:0009987	cellular process	603	-0.224238864	1.00E-04
GO:0034641	cellular nitrogen compound metabolic process	270	-0.285885568	0.00010001
GO:0006725	cellular aromatic compound metabolic process	222	-0.280367197	0.00010005

Within the cellular component category, the most predominant gene ontology terms were the GO:0032991 (protein-containing complex), GO:0018995 (host cellular component), GO:0043657 (host cell), GO:0030430 (host cell cytoplasm), GO:0033646 (host intracellular part), GO:0033643 (host cell part), and GO:0043656 (host intracellular region) with set size greater than 250 genes (Table 4.11).

Table 4.11. Top 10 enriched cellular component gene ontology terms in asymptomatic vs. naïve malaria.

ID	Ontology	Description	Set size	*p*-value
GO:0032991	CC	protein-containing complex	378	1.00E-04
GO:0018995	CC	host cellular component	339	1.00E-04
GO:0043657	CC	host cell	339	1.00E-04
GO:0030430	CC	host cell cytoplasm	264	0.0001
GO:0033646	CC	host intracellular part	264	0.0001
GO:0033643	CC	host cell part	289	0.0001
GO:0043656	CC	host intracellular region	265	0.0001
GO:0031982	CC	Vesicle	172	0.0001
GO:0020036	CC	Maurer's cleft	163	0.0001
GO:0005829	CC	Cytosol	151	0.0001

Upregulated genes in ASM versus UM were predominantly associated with cell adhesion, modulation by symbiont of host erythrocyte aggregation, modulation by symbiont of host cellular process, homotypic cell-cell adhesion, erythrocyte aggregation, evasion of host immune response, antigenic variation, response to biotic stimulus, and response to external biotic stimulus, (Figure 4.5). Conversely, downregulated genes in ASM versus UM were associated with protein-containing complex, host cellular component, host cell, and cellular process. In addition to the molecular function, several biological processes related to the metabolic process were suppressed, including primary metabolic processes, cellular metabolic processes, macromolecule metabolic processes, nitrogenous compound metabolic processes, and organic substance metabolic processes.

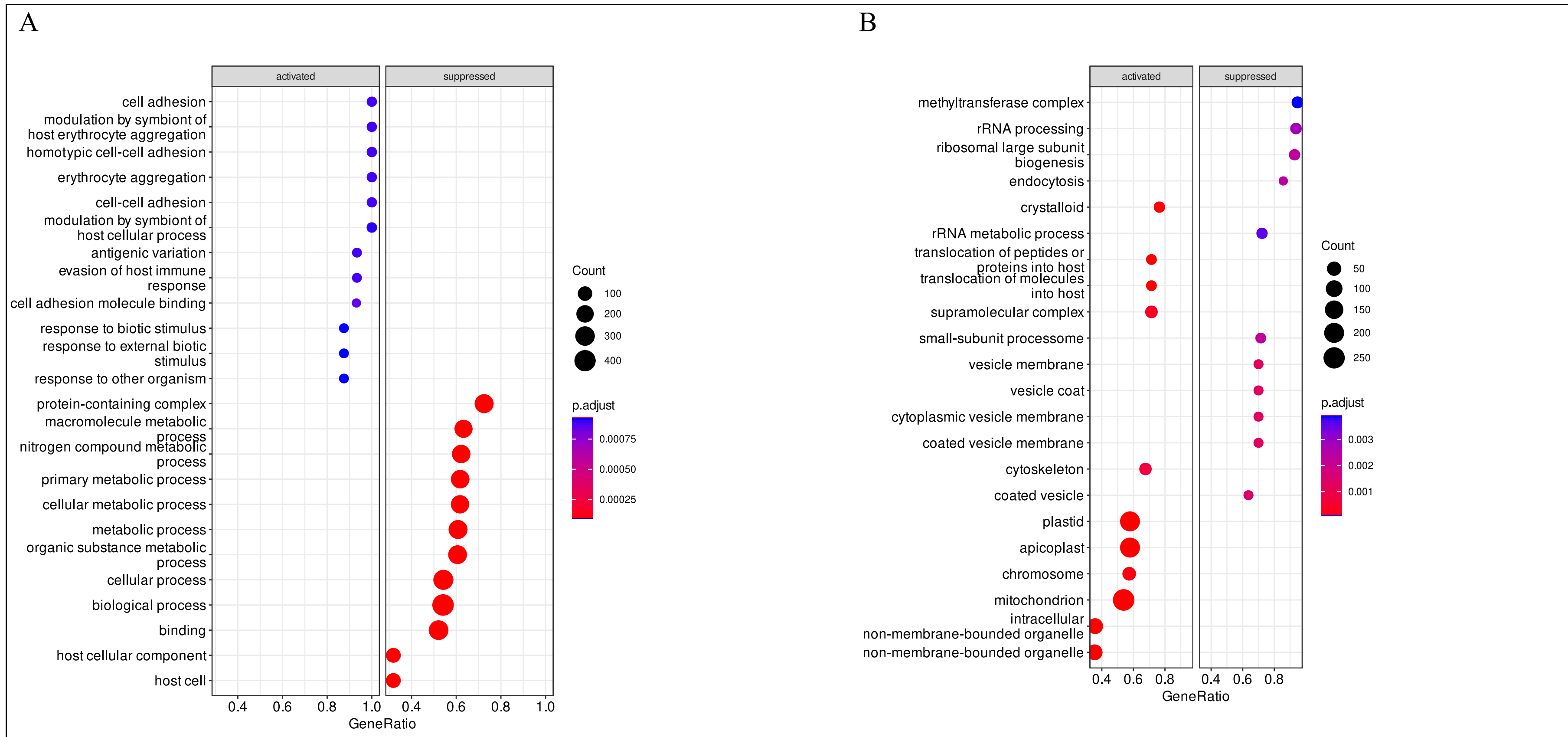


Figure 4.5. Scatterplots depicting enriched Gene Ontology terms for (a) asymptomatic compared to uncomplicated malaria, and (b) asymptomatic versus naïve malaria infections. The size of each dot represents its expression level, and the vertical bar denotes the significance level based on the adjusted *p*-value (*p-adjust*). Darker shades of red signify higher levels of significance.

Twenty-eight immune evasion genes associated with cell adhesion, modulation by symbiont of host cellular process, homotypic cell-cell adhesion, erythrocyte aggregation, antigenic variation, modulation by symbiont of host erythrocyte aggregation, evasion of host immune response, response to host immune response adhesion, response to biotic stimulus, response to other organisms, response to defenses of other organisms, response to host defenses, and response to external biotic stimulus were identified (Table 4.12). All these genes were identified as *Pf*EMP1 except one *rif* gene coding for the RIFIN protein

4.5.4 Functional Enrichment of the DEGs in ASM vs. UM and ASM vs. NM.

Using the KEGG pathways classification, enriched pathways and several suppressed pathways in asymptomatic vs. uncomplicated malaria were identified. Three KEGG pathways, malaria, lipoic acid metabolism, and DNA replication pathways, were activated out of the 14 top most enriched pathways (Figure 4.6). Ribosome and spliceosome pathways were largely suppressed in asymptomatic malaria compared to uncomplicated malaria.

Table 4.12. List of genes implicated in evasion of host immune response

Gene ID	Description
PF3D7_0937800	erythrocyte membrane protein 1
PF3D7_1240400	erythrocyte membrane protein 1
PF3D7_1041300	erythrocyte membrane protein 1
PF3D7_1100100	erythrocyte membrane protein 1
PF3D7_0632800	erythrocyte membrane protein 1
PF3D7_0632500	erythrocyte membrane protein 1
PF3D7_1200400	erythrocyte membrane protein 1
PF3D7_1240300	erythrocyte membrane protein 1
PF3D7_1100200	erythrocyte membrane protein 1
PF3D7_0808700	erythrocyte membrane protein 1
PF3D7_1000100	erythrocyte membrane protein 1
PF3D7_0400100	erythrocyte membrane protein 1
PF3D7_0733000	erythrocyte membrane protein 1
PF3D7_0412700	erythrocyte membrane protein 1
PF3D7_1200100	erythrocyte membrane protein 1
PF3D7_0800100	erythrocyte membrane protein 1
PF3D7_0426000	erythrocyte membrane protein 1
PF3D7_0800300	erythrocyte membrane protein 1
PF3D7_0421300	erythrocyte membrane protein 1
PF3D7_1219300	erythrocyte membrane protein 1
PF3D7_0223500	erythrocyte membrane protein 1
PF3D7_0420700	erythrocyte membrane protein 1
PF3D7_1255200	erythrocyte membrane protein 1
PF3D7_0800200	erythrocyte membrane protein 1
PF3D7_0808600	erythrocyte membrane protein 1
PF3D7_0712400	erythrocyte membrane protein 1
PF3D7_0100100	erythrocyte membrane protein 1
PF3D7_1000400	Rifin

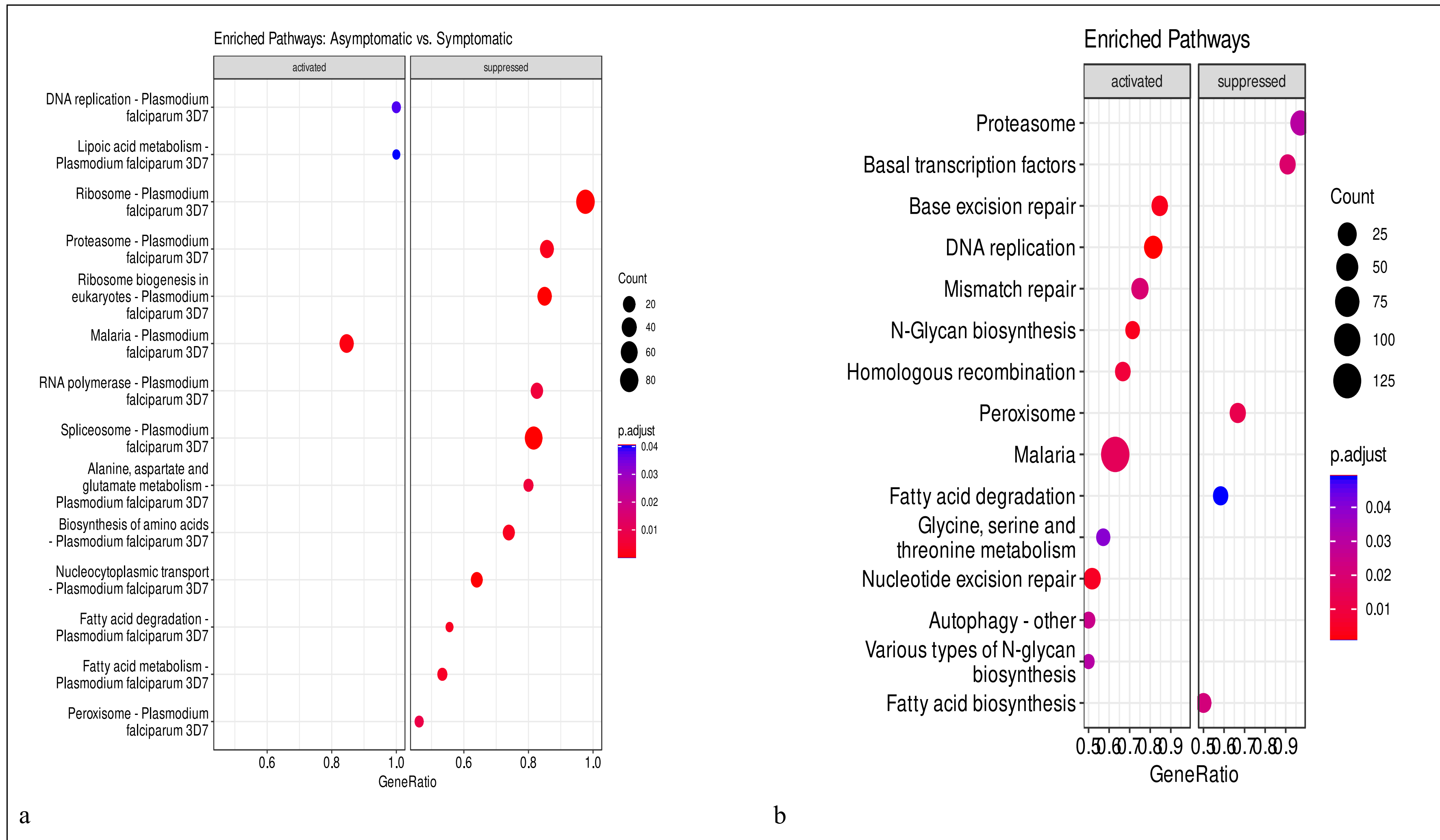


Figure 4.6 Scatter plot illustrating the top 14 Enriched KEGG Pathways, ranked by gene count and significance (-log10(*P*-value)). Panel (a) displays the enriched pathways in asymptomatic compared to uncomplicated malaria, while panel (b) shows the enriched pathways in asymptomatic compared to Naïve malaria. Each circle's color represents the adjusted *p*-value of the enriched pathway.

4.5.5 Differential Expression of Long Non-coding RNA in Asymptomatic Malaria Compared to Uncomplicated Malaria

An intersection of three coding predicting tools identified 1477 candidate lncRNAs. The fasta sequences of the identified lncRNAs can be accessed from https://github.com/JGisaina/supplimentary_files. To uncover lncRNAs involved in asymptomatic malaria infections, 23 samples from ASM (n = 11) and UM (n = 12) were analyzed. The principal component analysis revealed distinct clustering of these samples with 45.33% variation. PC1 accounted for 33.24% while PC2 accounted for 12.9% of the total variation, respectively (Figure 4.7)

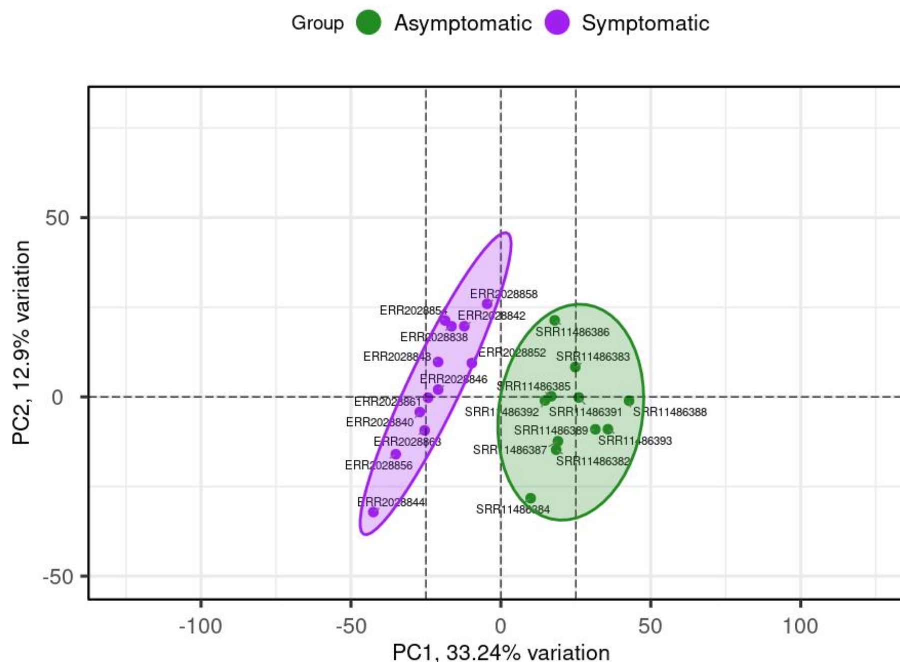


Figure 4.7. Cluster analysis of 23 (ASM) vs 12 (UM) samples by PCA. PC1 accounted for 32.2%, and PC2 accounted for 12.9% of the total variation.

Differential expression analysis identified 237 differentially expressed long noncoding RNA at the FDR<0.05, and a foldchange value of ± 2 relative to the normalized count's *p*-value was identified. Of these lncRNA, 216 (18%) were upregulated, and 21 (1.5%) were downregulated. The proportion of DE lncRNAs was higher compared to the down-regulated DE lncRNAs. The heatmap and volcano plot (Figure 4.8) illustrate the expression and clustering analysis of the lncRNAs.

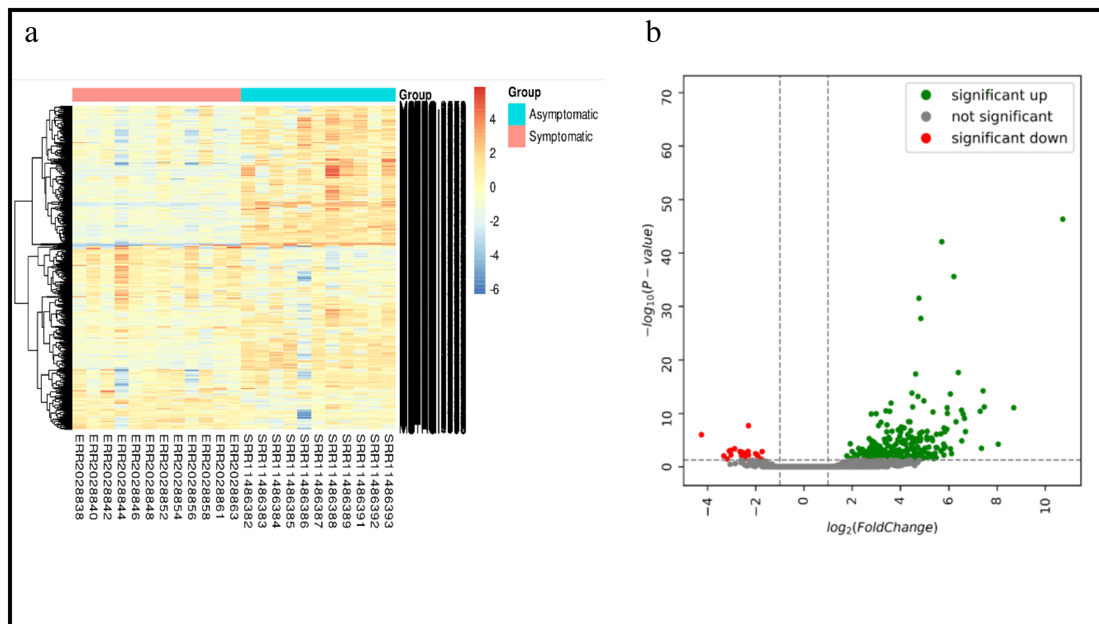


Figure 4.8. Expression Pattern Analysis. (a) Hierarchical clustering of differentially expressed lncRNA transcripts. Each row in the heat map corresponds to a distinct annotated gene, and the top horizontal bar delineates the various experimental groups under investigation. The color gradient in the vertical bars indicates the expression levels of the genes, ranging from non-detected (yellow) to upregulated (red) and downregulated (blue). (b) Volcano plot depicting results of the DESeq analysis. The horizontal axis indicates the log2 mean expression level, while the vertical axis shows the log2 transformed fold-change. The green dots represent genes that exhibited a notable increase in expression, while the red dots indicate genes that showed a decrease in expression. Grey dots represent genes that did not show significant differential expression.

4.5.6 Analysis of the Interaction of lncRNAs with Host Immune Genes

The possible function of the upregulated lncRNAs in the immunological process during asymptomatic infections was further investigated by analysing the interaction of the top 20 upregulated lncRNAs with *P. falciparum* immune evasion genes. The analysis identified five lncRNAs: MSTRG.1673, MSTRG.3921, MSTRG.4325, MSTRG.829, MSTRG.3083 (labelled as per StringTie unique IDs) that target six immune evasion genes (Table 4.13). The interactions mostly occurred in the upstream and the downstream regions of the genes.

Table 4.13. The interaction of upregulated long noncoding RNAs with *P. falciparum* genes implicated in immune evasion.

Query	Target gene	Target length	dG	ndG	Start Position Target	End Position Target
MSTRG.1673	PF3D7_0808600	8893	-9.47	-0.94	1	72
MSTRG.3921	PF3D7_0808600	8893	-12.18	-0.13	8800	8893
MSTRG.4325	PF3D7_0400100	8972	-4.05	-0.10	8895	8972
MSTRG.4325	PF3D7_1000400	1318	-2.47	-1.24	1308	1318
MSTRG.4325	PF3D7_0808600	8893	-10.86	-0.28	1	40
MSTRG.829	PF3D7_1041300	7548	-4.83	-0.12	7433	7548
MSTRG.829	PF3D7_0400100	8972	-3.31	-0.41	1	44
MSTRG.829	PF3D7_1240300	8938	-3.31	-0.41	1	44
MSTRG.829	PF3D7_0420700	6845	-4.38	-0.10	6797	6845
MSTRG.3083	PF3D7_1000400	1318	-4.54	-0.11	1	60
MSTRG.3083	PF3D7_0808600	8893	-5.97	-0.12	1	73

4.6 Expression of the *var* Genes Subsets in Asymptomatic Malaria Compared to Uncomplicated Malaria

4.6.1 Expression of *var* Genes Subsets in Asymptomatic Malaria Compared to Uncomplicated Malaria

No substantial variances were observed in the expression of *var* groups. However, seven enriched *var* genes, five downregulated, and two upregulated were identified (Figure 4.9). Five (CUFF.42, CUFF.50, CUFF.75, CUFF.219, and CUFF 203) enriched transcripts were identified as novel transcripts by Cufflinks. Notably, the downregulation of the *Plasmodium* Histone 3 (PF3D7_0610400) implicated in epigenetic regulation of the *var* genes expression was identified.

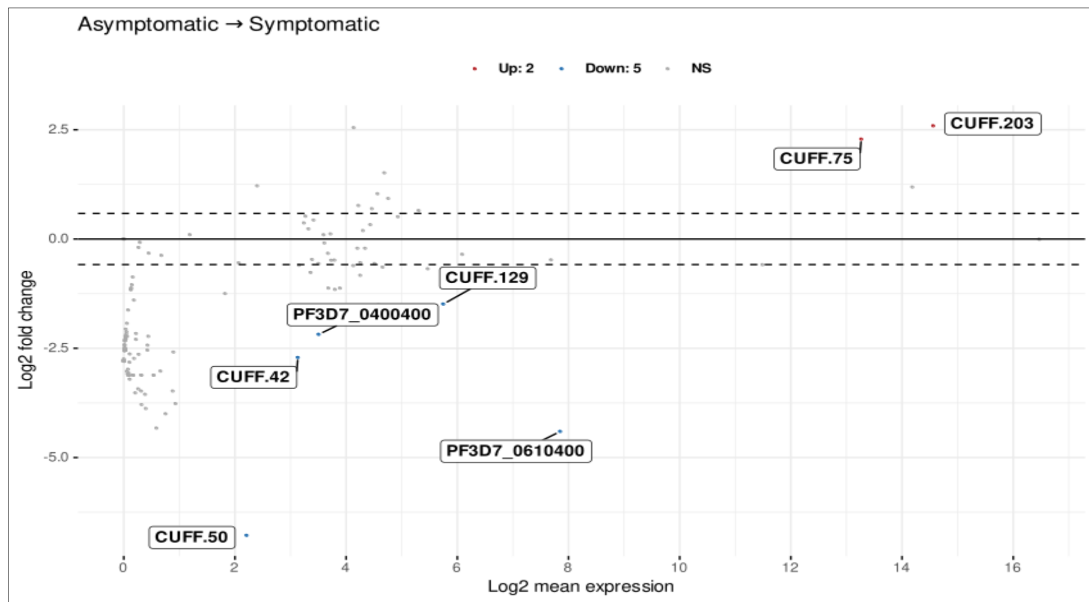


Figure 4.9. An MA plot showing the significantly enriched genes in asymptomatic malaria compared to uncomplicated malaria at FDR < 0.05 and log2FC values of ±1.5.

4.7 Single Nucleotide Polymorphisms within *var* Genes Associated with Asymptomatic Malaria

Variant calling identified 5714 raw variants. These variants were filtered to remove indels, resulting in 4198 SNPs. These SNPs were further filtered to remove poor-quality SNPs. Twenty high-quality SNPs were retained based on a quality score (QUAL) of greater than 30, Depth (DP) > 100, MAF > 0.05, and maximum missing rate of less than 20% (Table 4.14). Moreover, these SNPs were distributed across four chromosomes (Table 4.15).

Table 4.14. A table showing the 20 SNPs identified through variant calling in the *P. falciparum* virulent genes.

Chromosome	Position	Reference Allele	Alternate allele
NC_037280.1	103494	T	C
NC_037280.1	103718	A	T
NC_037280.1	103950	G	A
NC_037280.1	103969	T	C
NC_037280.1	104481	T	C
NC_037280.1	104982	G	A
NC_037280.1	105037	C	T
NC_004326.2	69585	T	C
NC_004326.2	69901	G	C
NC_004326.2	69961	G	T
NC_004326.2	69973	G	A
NC_004326.2	70045	A	G
NC_004326.2	70051	C	A
NC_004329.3	1366783	G	T
NC_004330.2	1420719	T	G
NC_004330.2	1420960	T	A
NC_004330.2	1420972	C	T
NC_004330.2	1421332	T	C
NC_004330.2	1422363	C	T
NC_004330.2	1422371	A	C

Table 4.15. The distribution of the 20 single nucleotide polymorphisms across the chromosomes

Chromosome	Length	Variants
NC_004326.2	1,343,557	6
NC_004329.3	1,472,805	1
NC_004330.2	1,541,735	6
NC_037280.1	947,102	7
Total	5,305,199	20

The impact of the 20 SNPs was predicted using the snpEff software, and the analysis identified 86 effects classified as low impact (4 effects =4.561%), moderate impact 16 effects = 18.605%), and modifier impact (66 effects =76.744%). Among the 20 SNPs, 16 (80%) had a missense effect, while 4 (20%) had a silent effect. The effects were further classified by region as downstream (30 effects =34.89%), exon (20 effects = 23.23%), transcript (20 effects = 23.23%), and upstream region (16 effects = 18.60%) (Figure 4.10).

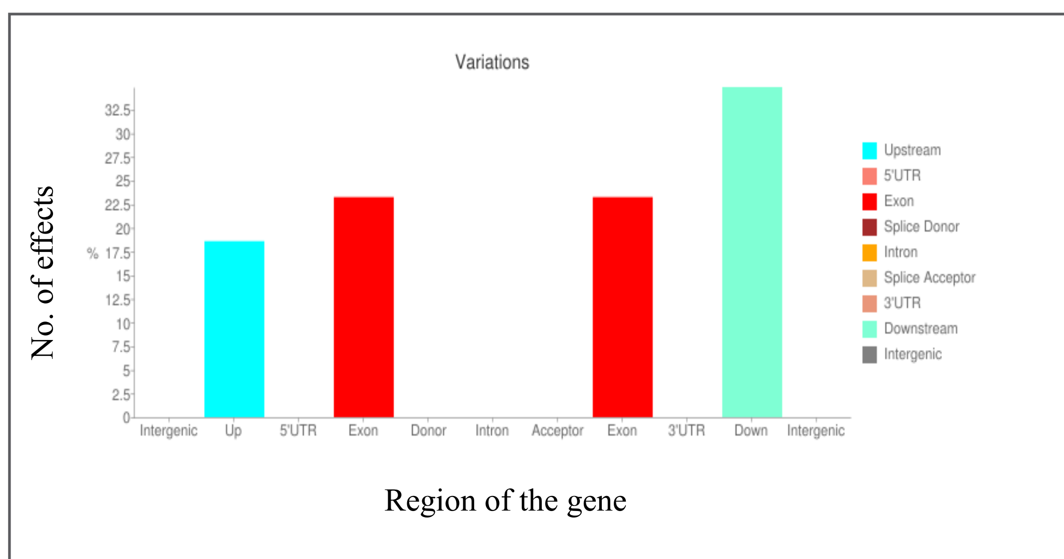


Figure 4.10. A bar plot showing the number of effects by region as either downstream, exon, or upstream region of a gene.

CHAPTER FIVE

DISCUSSIONS, CONCLUSIONS AND RECOMMENDATIONS

5.1 Discussion

Whereas studies have investigated the transcription signatures of *P. falciparum* across various forms of malaria, its transcriptional signature during asymptomatic malaria infections is underreported (Milner Jr et al., 2012; Lee et al., 2018). Furthermore, existing research has primarily focused on the host immunological response and the strategies employed by the parasite to evade the immune system in clinical cases of malaria (Bediako et al., 2019; Nallandhighal et al., 2019). Nonetheless, one study that aimed to understand the molecular factors contributing to the heterogeneity of different clinical malaria forms found significant differences in expression signatures between isolates from asymptomatic carriers and the *P. falciparum* 3D7-Lib strain (Almelli et al., 2014). Expanding on these findings, this study revealed significant differences in the transcription profiles of asymptomatic malaria compared to uncomplicated and naïve malaria. Remarkably, most variations in gene expression were noted in genes coding for proteins of unknown functions, as well as immune evasion genes like *var*, *phist*, and *rif* genes.

The enrichment analysis revealed an activation of pathways associated with cell adhesion, modulation of host erythrocyte aggregation by symbionts, response to biotic stimulus, homotypic cell-cell adhesion, erythrocyte aggregation, and response to host defenses (Figure 4.5). These results align with the observations made by Almeli et al. (2014), where they also noted an upregulation of pathways related to erythrocyte aggregation and adhesion to host cells in asymptomatic individuals compared to those in the 3D7-lib strain. Notably, the expression of genes encoding uncharacterized proteins underscores need for further functional annotation of *P. falciparum* genes.

Moreover, the transcriptional signature revealed 28 immune evasion genes (Table 4.12) that were identified as the factors that might be contributing to establishment of asymptomatic malaria infections. These genes were associated with host-parasite interaction for instance homotypic cell-cell adhesion, the erythrocyte aggregation by symbionts, response to biotic stimulus, and response to host defenses (Figure 4.5).

Interestingly, out of the 28 genes, 27 were virulent genes and one *rif* gene (PF3D7_1000400). The *rif* genes code for the RIFIN, a 27-45 kDa protein (Joannin et al., 2008). Studies have indicated that RIFINs could potentially aid in evading the immune system by interacting with human inhibitory immune receptors, although their specific functions are not yet fully understood (Saito et al., 2017; Sakoguchi et al., 2021). The central dogma of malaria pathogenesis involves variability of antigens and how the parasite evades the host immune system. These mechanisms are mediated by variant surface antigens (VSAs), particularly the *Pf*EMP1 antigens (Roberts et al., 1992). The mutual expression and switching off of the *var* genes could be mechanisms for chronic symptomatic malaria infections.

Upregulation of a subset of CVGs showing hypervariability, including *stevor*, *rifin*, and *pfmc-2tm,* and downregulation of a subset of *phist* genes was observed. The clonal variant genes have been grouped into those showing hypervariability between *Plasmodium* isolates and which are driven by immune evasion (*rif, stevor, var, pfmc-2tm*) and those conserved across *Plasmodium* isolates that are implicated in the functional diversification of the parasites (including; *fikk, extramp, phist, dnaj,* and *cs*, among others (Templeton, 2009; Rovira-Graells et al., 2012). Furthermore, another investigation into the expression of both groups of CVGs indicates that variations in transcription within *Plasmodium* parasites lead to differences in both antigenic properties and functionality (Rovira-Graells et al., 2012). However, the findings of this study suggests that transcriptional variation of CVGs during asymptomatic infection is driven by immune evasion rather than functional diversification.

The intraerythrocytic developmental cycle (IDC) plays a crucial role in *P. falciparum*'s life cycle. This 48-hour cycle is characterized by rapid growth through repeated cycles of infection and asexual replication, leading to the emergence of clinical symptoms and an increase in parasite biomass during the symptomatic phase of infection (Miller et al., 2002). Carbon metabolism serves as the primary energy source during this developmental stage. A study investigating the transcriptome profile of the IDC discovered increased expression of various genes associated with carbon metabolism (Bozdech et al., 2003). Unlike in symptomatic infections, the study findings indicate suppression of various metabolic pathways related to carbon metabolism, fatty acid biosynthesis, and overall metabolism during asymptomatic malaria infections. These

findings suggest that *P. falciparum* employs a strategy of suppressing central-catabolic metabolism to maintain low parasitemia during asymptomatic infection. Research into *P. falciparum* metabolism under isoleucine-starved conditions has shown that the parasite adapts by suppressing central-carbon metabolism intermediates (Babbitt et al., 2012). Since the same metabolic pathways are suppressed during asymptomatic infections, it is suggested that the parasite may adjust its metabolism similarly to establish asymptomatic infections, although further investigation is necessary to confirm this hypothesis.

Long non-coding RNAs (lncRNAs) interact with mRNA, influencing various biological processes such as signaling pathways, RNA splicing, and DNA regulation (Ayub et al., 2019; Ahmad et al., 2021). In *P. falciparum*, these RNAs play roles in regulating immunological responses and signaling pathways through epigenetic mechanisms (Chen et al., 2022). However, their specific functions in asymptomatic malaria infections are not well comprehended. The current study revealed interaction of lncRNAs with *P. falciparum* genes implicated in immune response genes and host-parasite interaction. Five lncRNAs found to interact with six *var* genes are implicated in *Plasmodium* immune evasion (Table 4.13). These interactions occurred mainly in the upstream regions of the implicated genes, which could be promoter-binding sites. Since it has been shown that lncRNAs can directly interact with the miRNA or an RNA protein binding site in the target mRNA and initiate regulatory property (Wanowska et al., 2018; Sebastian-delaCruz et al., 2021), these findings suggest that the lncRNAs could be using similar mechanisms to regulate the expression of implicated mRNAs during asymptomatic infections. However, in-depth immunoprecipitation assays are required to validate these findings.

Since specific group of *var* genes are linked to various types of clinical malaria (Bertin et al., 2013), this study investigated if there exists a distinct set of genes that are associated with asymptomatic infections by examining their transcription patterns in ASM vs. UM. Additionally, the activity of regulatory genes involved in *var* switching was evaluated to comprehend their impact on *var* gene expression between asymptomatic carriers and individuals with uncomplicated malaria. The analysis revealed no significant alterations in the expression of the *var* gene groups. However, differential expression in a few uncharacterized transcripts was observed. (CUFF.42,

CUFF.50, CFF.75, CUFF.129, and CUFF.203). Moreover, this study noted downregulation of the *P. falciparum* putative nuclear pore protein nucleoporin, NUP116/NSP116 (PF3D7_1473700), and the Histone 3 (H3) factors that regulate *P. falciparum var* switching. The process of *var* switching is not primarily dependent on transcriptional factors such as cis-acting elements but rather epigenetic regulation, which relies on heterochromatin remodeling and localization of a locus within the nucleus. This transcription occurs at an undefined perinuclear site away from the 59 repressed gene clusters (Ralph et al., 2005). The nucleoporin plays a significant role in chromatin organization and transcriptional regulation of genes (Capelson et al., 2010). A study of perinuclear PfNup116 in *var* monoallelic expression using PfNup116 antibodies concluded that the expression is independent of the perinuclear pore (Guizetti et al., 2013); consequently, the precise function of Nup 116 in gene regulation remains unclear.

Accurate detection and quantification of asymptomatic malaria infections, whether microscopic or sub-microscopic, is crucial for transmission-blocking and community intervention programs (Bousema et al., 2014). New treatment strategies that target the asymptomatic malaria reservoir, such as mass drug administration, open new avenues to tackle malaria. However, low-density asymptomatic infections are poorly detected by commercially available rapid diagnostic tools (mRDTs) as well as standard microscopy (Hofer et al., 2023). Hence, effective detection of asymptomatic infections largely relies on quantitative real-time and nested PCR (Agbana et al., 2022). The findings of this study indicate that two transcripts, PF3D7_0930000 (putative procollagen lysine 5-dioxygenase protein) and PF3D7_1471700 (protein with unknown function) are upregulated in cases of asymptomatic malaria infections. These transcripts rank among the top 10 upregulated in both comparisons and have been previously identified as gametocyte-enriched transcripts (Essuman et al., 2017). They serve as potential biomarkers for developing better molecular diagnostic tools for accurate detection and treatment of asymptomatic infections. This is crucial because current mRDTs, which rely on *P. falciparum* histidine-rich protein (HRP2), are encountering notable drawbacks due to the reported gene deletions in some parasites (Prosser et al., 2021).

Several SNPs conferring fitness advantage traits to *P. falciparum* in drug resistance

have been identified. It is unknown whether there is a single nucleotide polymorphism that influences immune response during asymptomatic malaria infection. The genomic analysis identified single nucleotide polymorphism in two *Plasmodium* erythrocyte coding genes, namely PF3D7_0800100 and PF3D7_0900100. These genes were among the significantly downregulated in asymptomatic vs. uncomplicated malaria infections. The study did not find alleles that were exclusively associated with asymptomatic malaria. However, one mutation occurring in the PF3D7_0900100 gene at chromosome NC_004330.2 position 1422371, which was occurring in samples isolated from uncomplicated malaria samples (Table 4.13). This gene was downregulated in asymptomatic malaria, and its function in asymptomatic malaria infections is not clear.

5.2 Conclusions

This study established that:

1. Antigenic variation, adaptation through clonally variant gene expression, and suppression of the carbon metabolism are adaptive mechanisms for *P. falciparum* in establishing asymptomatic malaria infections.
2. Transcriptional variability of clonal variant genes mediate *P. falciparum* adaption to changing environments beyond immune evasion.
3. The observed large transcriptional changes occurring in transcripts encoding proteins of unknown function suggests their significance in establishing asymptomatic infections.
4. *Plasmodium falciparum* does not express specific group of its virulent genes during asymptomatic malaria infections.
5. There are no specific SNPs in the virulent genes that are associated with asymptomatic malaria infections.

5.3 Limitations

1. MiSeq sequencing platform was used in this study. Other advanced platforms exist and are able to generate larger number of reads.
2. The retrieved datasets were generated from different time points and, therefore, subjected to batch effect.

3. This study did not report host transcription patterns during asymptomatic infection.
4. Validation of *var* gene expression by use of qPCR was not done
5. While the parasite may largely be using fitness advantage to remain undetected in asymptomatic infections, the study did not exhaustively examine the role of host immunological factors that could maximize parasite fitness advantage and influence the asymptomatic phenotype.

5.4 Recommendations

1. Dual sequencing analysis is recommended to simultaneously elucidate host-parasite interactions and genetic factors conferring fitness advantage to *P. falciparum* during asymptomatic infections.
2. There is need for functional annotation of the genes coding for proteins of unknown functions to uncover their role in asymptomatic infections
3. Evaluation of PF3D7_0930000 and PF3D7_1471700 transcripts as potential biomarkers for detection of asymptomatic malaria infections
4. Validating the observed gene expression profiles at a protein level will also be important.

REFERENCES

Acquah, F. K., Donu, D., Bredu, D., Eyia-Ampah, S., Amponsah, J. A., Quartey, J., Obboh, E. K., Mawuli, B. A., & Amoah, L. E. (2020). Asymptomatic carriage of Plasmodium falciparum by individuals with variant blood groups and haemoglobin genotypes in southern Ghana. *Malaria journal*, *19*, 1-8.

Aderem, A., Adkins, J. N., Ansong, C., Galagan, J., Kaiser, S., Korth, M. J., Law, G. L., McDermott, J. G., Proll, S. C., & Rosenberger, C. (2011). A systems biology approach to infectious disease research: innovating the pathogen-host research paradigm. *Mbio*, *2*(1), 10.1128/mbio. 00325-00310.

Agbana, H. B., Rogier, E., Lo, A., Abukari, Z., Jones, S., Gyan, B., Aidoo, M., & Amoah, L. E. (2022). Detecting asymptomatic carriage of Plasmodium falciparum in southern Ghana: utility of molecular and serological diagnostic tools. *Malaria journal*, *21*(1), 57.

Almelli, T., Nuel, G., Bischoff, E., Aubouy, A., Elati, M., Wang, C. W., Dillies, M.-A., Coppée, J.-Y., Ayissi, G. N., & Basco, L. K. (2014). Differences in gene transcriptomic pattern of Plasmodium falciparum in children with cerebral malaria and asymptomatic carriers. *PLoS One*, *9*(12), e114401.

Amit-Avraham, I., Pozner, G., Eshar, S., Fastman, Y., Kolevzon, N., Yavin, E., & Dzikowski, R. (2015). Antisense long noncoding RNAs regulate var gene activation in the malaria parasite Plasmodium falciparum. *Proceedings of the National Academy of Sciences*, *112*(9), E982-E991.

Andagalu, B., Watson, O. J., Onyango, I., Opot, B., Okoth, R., Chemwor, G., Sifuna, P., Juma, D., Cheruiyot, A., & Yeda, R. (2023). Malaria Transmission Dynamics in a High-Transmission Setting of Western Kenya and the Inadequate Treatment Response to Artemether-Lumefantrine in an Asymptomatic Population. *Clinical Infectious Diseases*, *76*(4), 704-712.

Ariey, F., Hommel, D., Le Scanf, C., Duchemin, J. B., Peneau, C., Hulin, A., Sarthou, J. L., Reynes, J. M., Fandeur, T., & Mercereau-Puijalon, O. (2001). Association of severe malaria with a specific Plasmodium falciparum genotype in French Guiana. *The Journal of infectious diseases*, *184*(2), 237-241.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., & Eppig, J. T. (2000). Gene ontology: tool for the unification of biology. *Nature genetics*, *25*(1), 25-29.

Babbitt, S. E., Altenhofen, L., Cobbold, S. A., Istvan, E. S., Fennell, C., Doerig, C., Llinás, M., & Goldberg, D. E. (2012). Plasmodium falciparum responds to amino acid starvation by entering into a hibernatory state. *Proceedings of the National Academy of Sciences*, *109*(47), E3278-E3287.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, *116*(2), 281-297.

Bartoloni, A., & Zammarchi, L. (2012). Clinical aspects of uncomplicated and severe malaria. *Mediterranean journal of hematology and infectious diseases*, *4*(1).

Bediako, Y., Adams, R., Reid, A. J., Valletta, J. J., Ndungu, F. M., Sodenkamp, J., Mwacharo, J., Ngoi, J. M., Kimani, D., & Kai, O. (2019). Repeated clinical malaria episodes are associated with modification of the immune system in children. *BMC medicine*, *17*, 1-14.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerfµl approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289-300.

Bernabeu, M., Danziger, S. A., Avril, M., Vaz, M., Babar, P. H., Brazier, A. J., Herricks, T., Maki, J. N., Pereira, L., & Mascarenhas, A. (2016). Severe adult malaria is associated with specific PfEMP1 adhesion types and high parasite biomass. *Proceedings of the National Academy of Sciences*, *113*(23), E3270-E3279.

Bertin, G. I., Lavstsen, T., Guillonneau, F., Doritchamou, J., Wang, C. W., Jespersen, J. S., Ezimegnon, S., Fievet, N., Alao, M. J., & Lalya, F. (2013). Expression of the domain cassette 8 Plasmodium falciparum erythrocyte membrane protein 1 is associated with cerebral malaria in Benin. *PLoS One*, *8*(7), e68368.

Boldt, A. B., Van Tong, H., Grobusch, M. P., Kalmbach, Y., Ella, A. D., Kombila, M., Meyer, C. G., Kun, J. F., Kremsner, P. G., & Velavan, T. P. (2019). The blood transcriptome of childhood malaria. *EBioMedicine*, *40*, 614-625.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120.

Bousema, T., Okell, L., Felger, I., & Drakeley, C. (2014). Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nature Reviews Microbiology*, *12*(12), 833-840.

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.

Bowen, A. J., & Corcoran, A. E. (2008). How chromatin remodelling allows shuffling of immunoglobulin heavy chain genes. *Molecular BioSystems*, *4*(8), 790-798.

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.

Broadbent, K. M., Broadbent, J. C., Ribacke, U., Wirth, D., Rinn, J. L., & Sabeti, P. C. (2015). Strand-specific RNA sequencing in Plasmodium falciparum malaria identifies developmentally regulated long non-coding RNA and circular RNA. *BMC genomics*, *16*(1), 1-22.

Bull, P. C., Buckee, C. O., Kyes, S., Kortok, M. M., Thathy, V., Guyah, B., Stoute, J. A., Newbold, C. I., & Marsh, K. (2008). Plasmodium falciparum antigenic variation. Mapping mosaic var gene sequences onto a network of shared, highly polymorphic sequence blocks. *Molecular microbiology*, *68*(6), 1519-1534.

Butler, N. S., Harris, T. H., & Blader, I. J. (2013). Regulation of immunopathogenesis during Plasmodium and Toxoplasma infections: more parallels than distinctions? *Trends in parasitology*, *29*(12), 593-602.

Cabantous, S., Doumbo, O., Poudiougou, B., Louis, L., Barry, A., Oumar, A. A., Traore, A., Marquet, S., & Dessein, A. (2017). Gene expression analysis reveals genes common to cerebral malaria and neurodegenerative disorders. *The Journal of infectious diseases*, *216*(6), 771-775.

Cheaveau, J., Mogollon, D. C., Mohon, M. A. N., Golassa, L., Yewhalaw, D., & Pillai, D. R. (2019). Asymptomatic malaria in the clinical and public health context. *Expert review of anti-infective therapy*, *17*(12), 997-1010.

Chen, G., Liu, S.-c., Fan, X.-y., Jin, Y.-l., Li, X., & Du, Y.-t. (2022). Plasmodium manipulates the expression of host long non-coding RNA during red blood cell intracellular infection. *Parasites & vectors*, *15*(1), 1-15.

Chen, I., Clarke, S. E., Gosling, R., Hamainza, B., Killeen, G., Magill, A., O'Meara, W., Price, R. N., & Riley, E. M. (2016). "Asymptomatic" malaria: a chronic and debilitating infection that should be treated. *PLoS medicine*, *13*(1), e1001942.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *fly*, *6*(2), 80-92.

Consortium, G. O. (2019). The gene ontology resource: 20 years and still GOing strong. *Nucleic acids research*, *47*(D1), D330-D338.

Cortés, A., & Deitsch, K. W. (2017). Malaria epigenetics. *Cold Spring Harbor Perspectives in Medicine*, *7*(7).

Cowell, A., & Winzeler, E. (2018). Exploration of the Plasmodium falciparum resistome and druggable genome reveals new mechanisms of drug resistance and antimalarial targets. *Microbiology insights*, *11*, 1178636118808529.

Cox, F. E. (2010). History of the discovery of the malaria parasites and their vectors. *Parasites & vectors*, *3*(1), 1-9.

Dandewad, V., Vindu, A., Joseph, J., & Seshadri, V. (2019). Import of human miRNA-RISC complex into Plasmodium falciparum and regulation of the parasite gene expression. *Journal of biosciences*, *44*(2), 50.

Deitsch, K. W., & Dzikowski, R. (2017). Variant gene expression and antigenic variation by malaria parasites. *Annual review of microbiology*, *71*, 625-641.

Douglas, N. M., Pontororing, G. J., Lampah, D. A., Yeo, T. W., Kenangalem, E., Poespoprodjo, J. R., Ralph, A. P., Bangs, M. J., Sugiarto, P., & Anstey, N. M. (2014). Mortality attributable to Plasmodium vivaxmalaria: a clinical audit from Papua, Indonesia. *BMC medicine*, *12*(1), 1-13.

Epp, C., Li, F., Howitt, C. A., Chookajorn, T., & Deitsch, K. W. (2009). Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite Plasmodium falciparum. *Rna*, *15*(1), 116-127.

Essuman, E., Grabias, B., Verma, N., Chorazeczewski, J. K., Tripathi, A. K., Mlambo, G., Addison, E. A., Amoah, A. G., Quakyi, I., & Oakley, M. S. (2017). A novel gametocyte biomarker for superior molecular detection of the Plasmodium falciparum infectious reservoirs. *The Journal of infectious diseases*, *216*(10), 1264-1272.

Frimpong, A., Amponsah, J., Adjokatseh, A. S., Agyemang, D., Bentum-Ennin, L., Ofori, E. A., Kyei-Baafour, E., Akyea-Mensah, K., Adu, B., & Mensah, G. I. (2020). Asymptomatic malaria infection is maintained by a balanced pro-and anti-inflammatory response. *Frontiers in microbiology*, *11*, 559255.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., & Bowman, S. (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature*, *419*(6906), 498-511.

Gazzinelli, R. T., Kalantari, P., Fitzgerald, K. A., & Golenbock, D. T. (2014). Innate sensing of malaria parasites. *Nature Reviews Immunology*, *14*(11), 744-757.

Greenwood, J. M., Ezquerra, A. L., Behrens, S., Branca, A., & Mallet, L. (2016). Current analysis of host–parasite interactions with a focus on next generation sequencing data. *Zoology*, *119*(4), 298-306.

Griffin, J. T., Hollingsworth, T. D., Reyburn, H., Drakeley, C. J., Riley, E. M., & Ghani, A. C. (2015). Gradual acquisition of immunity to severe malaria with increasing exposure. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801), 20142657.

Grobusch, M., & Kremsner, P. (2005). Uncomplicated malaria. *Malaria: Drugs, disease and post-genomic biology*, 81-104.

Hanahan, D. (2022). Hallmarks of cancer: new dimensions. *Cancer discovery*, *12*(1), 31-46.

Harrison, P. W., Wright, A. E., & Mank, J. E. (2012). The evolution of gene expression and the transcriptome–phenotype relationship. Seminars in Cell & Developmental Biology, (Vol. 23, pp. 222-229).

Hodgson, S. H., Muller, J., Lockstone, H. E., Hill, A. V., Marsh, K., Draper, S. J., & Knight, J. C. (2019). Use of gene expression studies to investigate the human immunological response to malaria infection. *Malaria journal*, *18*, 1-15.

Hofer, L. M., Kweyamba, P. A., Sayi, R. M., Chabo, M. S., Maitra, S. L., Moore, S. J., & Tambwe, M. M. (2023). Malaria rapid diagnostic tests reliably detect asymptomatic Plasmodium falciparum infections in school-aged children that are infectious to mosquitoes. *Parasites & vectors*, *16*(1), 217.

Holliday, R. (1987). The inheritance of epigenetic defects. *Science*, *238*(4824), 163-170.

Howes, R. E., Battle, K. E., Mendis, K. N., Smith, D. L., Cibulskis, R. E., Baird, J. K., & Hay, S. I. (2016). Global epidemiology of Plasmodium vivax. *The American journal of tropical medicine and hygiene*, *95*(6 Suppl), 15.

Jain, P., Chakma, B., Patra, S., & Goswami, P. (2014). Potential biomarkers and their applications for rapid and reliable detection of malaria. *BioMed research international*, *2014*.

Joannin, N., Abhiman, S., Sonnhammer, E. L., & Wahlgren, M. (2008). Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. *BMC genomics*, *9*, 1-14.

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology*, *37*(8), 907-915.

Kimenyi, K. M., Wamae, K., & Ochola-Oyier, L. I. (2019). Understanding P. falciparum asymptomatic infections: a proposition for a transcriptomic approach. *Frontiers in immunology*, *10*, 2398.

Kimura, T. (2020). Non-coding natural antisense RNA: mechanisms of action in the regulation of target gene expression and its clinical implications. *Yakugaku zasshi: Journal of the Pharmaceutical Society of Japan*, *140*(5), 687-700.

Kojom Foko, L. P., Kumar, A., Hawadak, J., & Singh, V. (2023). Plasmodium cynomolgi in humans: current knowledge and future directions of an emerging zoonotic malaria parasite. *Infection*, *51*(3), 623-640.

Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., & Gao, G. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic acids research*, *35*(suppl_2), W345-W349.

Kraemer, S. M., & Smith, J. D. (2003). Evidence for the importance of genetic structuring to the structural and functional specialization of the Plasmodium falciparum var gene family. *Molecular microbiology*, *50*(5), 1527-1538.

Kyriacou, H. M., Stone, G. N., Challis, R. J., Raza, A., Lyke, K. E., Thera, M. A., Koné, A. K., Doumbo, O. K., Plowe, C. V., & Rowe, J. A. (2006). Differential var gene transcription in Plasmodium falciparum isolates from patients with cerebral malaria compared to hyperparasitaemia. *Molecular and biochemical parasitology*, *150*(2), 211-218.

Laishram, D. D., Sutton, P. L., Nanda, N., Sharma, V. L., Sobti, R. C., Carlton, J. M., & Joshi, H. (2012). The complexities of malaria disease manifestations with a focus on asymptomatic malaria. *Malaria journal*, *11*(1), 1-15.

Lamptey, H., Seidu, Z., Lopez-Perez, M., Kyei-Baafour, E., Hviid, L., Adjei, G. O., & Ofori, M. F. (2023). Impact of haemoglobinopathies on asymptomatic Plasmodium falciparum infection and naturally acquired immunity among children in Northern Ghana. *Frontiers in Hematology*, *2*, 1150134.

Lee, H. J., Georgiadou, A., Walther, M., Nwakanma, D., Stewart, L. B., Levin, M., Otto, T. D., Conway, D. J., Coin, L. J., & Cunnington, A. J. (2018). Integrated pathogen load and dual transcriptome analysis of systemic host-pathogen interactions in severe malaria. *Science translational medicine*, *10*(447), eaar3619.

Li, A., Zhang, J., & Zhou, Z. (2014). PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC bioinformatics*, *15*, 1-10.

Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*, *30*(20), 2843-2851.

Li, X., Peng, J., & Yi, C. (2021). The epitranscriptome of small non-coding RNAs. *Non-coding RNA Research*, *6*(4), 167-173.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923-930.

Lindblade, K. A., Steinhardt, L., Samuels, A., Kachur, S. P., & Slutsker, L. (2013). The silent threat: asymptomatic parasitemia and malaria transmission. *Expert review of anti-infective therapy*, *11*(6), 623-639.

Liu, L., Chen, Z., Yang, Y., Xiao, Y., Liu, C., Ma, Y., Soberón, M., Bravo, A., Yang, Y., & Liu, K. (2018). A single amino acid polymorphism in ABCC2 loop 1 is responsible for differential toxicity of Bacillus thuringiensis Cry1Ac toxin in different Spodoptera (Noctuidae) species. *Insect biochemistry and molecular biology*, *100*, 59-65.

Lodde, V., Floris, M., Muroni, M. R., Cucca, F., & Idda, M. L. (2022). Non-coding RNAs in malaria infection. *Wiley Interdisciplinary Reviews: RNA*, *13*(3), e1697.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and

dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*(12), 1-21.

Mace, K. E., Arguin, P. M., & Tan, K. R. (2018). Malaria surveillance—United States, 2015. *MMWR Surveillance Summaries*, *67*(7), 1.

Mackintosh, C. L., Beeson, J. G., & Marsh, K. (2004). Clinical features and pathogenesis of severe malaria. *Trends in parasitology*, *20*(12), 597-603.

Magesa, S., Mdira, K., Babiker, H., Alifrangis, M., Färnert, A., Simonsen, P., Bygbjerg, I., Walliker, D., & Jakobsen, P. (2002). Diversity of Plasmodium falciparum clones infecting children living in a holoendemic area in north-eastern Tanzania. *Acta Tropica*, *84*(2), 83-92.

Massah, S., Beischlag, T. V., & Prefontaine, G. G. (2015). Epigenetic events regulating monoallelic gene expression. *Critical reviews in biochemistry and molecular biology*, *50*(4), 337-358.

Merrick, C. J., Huttenhower, C., Buckee, C., Amambua-Ngwa, A., Gomez-Escobar, N., Walther, M., Conway, D. J., & Duraisingh, M. T. (2012). Epigenetic dysregulation of virulence gene expression in severe Plasmodium falciparum malaria. *The Journal of infectious diseases*, *205*(10), 1593-1600.

Miller, Baruch, D., Marsh, K., & Doumbo, O. (2002). The pathogenic basis of malaria. *Nature*, *415*(6872), 673-679.

Milner Jr, D. A., Pochet, N., Krupka, M., Williams, C., Seydel, K., Taylor, T. E., Van de Peer, Y., Regev, A., Wirth, D., & Daily, J. P. (2012). Transcriptional profiling of Plasmodium falciparum parasites from patients with severe malaria identifies distinct low vs. high parasitemic clusters. *PLoS One*, *7*(7), e40739.

Mooney, J. P., DonVito, S. M., Jahateh, M., Bittaye, H., Bottomley, C., D'Alessandro, U., & Riley, E. M. (2022). Dry season prevalence of Plasmodium falciparum in asymptomatic gambian children, with a comparative evaluation of diagnostic methods. *Malaria journal*, *21*(1), 171.

Nallandhighal, S., Park, G. S., Ho, Y.-Y., Opoka, R. O., John, C. C., & Tran, T. M. (2019). Whole-blood transcriptional signatures composed of erythropoietic and NRF2-regulated genes differ between cerebral malaria and severe malarial anemia. *The Journal of infectious diseases*, *219*(1), 154-164.

Nkumama, I. N., O'meara, W. P., & Osier, F. H. (2017). Changes in malaria epidemiology in Africa and new challenges for elimination. *Trends in parasitology*, *33*(2), 128-140.

Nyarko, P. B., & Claessens, A. (2021). Understanding host–pathogen–vector interactions with chronic asymptomatic malaria infections. *Trends in parasitology*, *37*(3), 195-204.

Park, Y. H., Shi, Y. P., Liang, B., Medriano, C. A. D., Jeon, Y. H., Torres, E., Uppal, K., Slutsker, L., & Jones, D. P. (2015). High-resolution metabolomics to discover potential parasite-specific biomarkers in a Plasmodium falciparum erythrocytic stage culture system. *Malaria journal*, *14*, 1-9.

Pertea, G., & Pertea, M. (2020). GFF utilities: GffRead and GffCompare. *F1000Research*, *9*, 304.

Portugal, S., Moebius, J., Skinner, J., Doumbo, S., Doumtabe, D., Kone, Y., Dia, S., Kanakabandi, K., Sturdevant, D. E., & Virtaneva, K. (2014). Exposure-dependent control of malaria-induced inflammation in children. *PLoS Pathogens*, *10*(4), e1004079.

Prosser, C., Gresty, K., Ellis, J., Meyer, W., Anderson, K., Lee, R., & Cheng, Q. (2021). Plasmodium falciparum histidine-rich protein 2 and 3 gene deletions in strains from Nigeria, Sudan, and South Sudan. Emerging infectious diseases, 27(2), 471.

Quinlan, A. R. (2014). BEDTools: the Swiss-army tool for genome feature analysis. *Current protocols in bioinformatics*, *47*(1), 11.12. 11-11.12. 34.

Robert, F., & Pelletier, J. (2018). Exploring the impact of single-nucleotide polymorphisms on translation. *Frontiers in genetics*, *9*, 507.

Roberts, D. J., Craig, Berendt, A. R., Pinches, R., Nash, G., Marsh, K., & Newbold, C. I. (1992). Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature*, *357*(6380), 689-692.

Rogerson, S. J., Desai, M., Mayor, A., Sicuri, E., Taylor, S. M., & van Eijk, A. M. (2018). Burden, pathology, and costs of malaria in pregnancy: new developments for an old problem. *The Lancet infectious diseases*, *18*(4), e107-e118.

Rottmann, M., Lavstsen, T., Mugasa, J. P., Kaestli, M., Jensen, A. T., Müller, D., Theander, T., & Beck, H.-P. (2006). Differential expression of var gene groups is associated with morbidity caused by Plasmodium falciparum infection in Tanzanian children. *Infection and immunity*, *74*(7), 3904-3911.

Rovira-Graells, N., Gupta, A. P., Planet, E., Crowley, V. M., Mok, S., De Pouplana, L. R., Preiser, P. R., Bozdech, Z., & Cortés, A. (2012). Transcriptional variation in the malaria parasite Plasmodium falciparum. *Genome research*, *22*(5), 925-938.

Saito, F., Hirayasu, K., Satoh, T., Wang, C. W., Lusingu, J., Arimori, T., Shida, K., Palacpac, N. M. Q., Itagaki, S., & Iwanaga, S. (2017). Immune evasion of Plasmodium falciparum by RIFIN via inhibitory receptors. *Nature*, *552*(7683), 101-105.

Sakoguchi, A., Saito, F., Hirayasu, K., Shida, K., Matsuoka, S., Itagaki, S., Nakai, W., Kohyama, M., Suenaga, T., & Iwanaga, S. (2021). Plasmodium falciparum RIFIN is a novel ligand for inhibitory immune receptor LILRB2. *Biochemical and biophysical research communications*, *548*, 167-173.

Saliminejad, K., Khorram Khorshid, H. R., Soleymani Fard, S., & Ghaffari, S. H. (2019). An overview of microRNAs: Biology, functions, therapeutics, and

analysis methods. *Journal of cellular physiology*, *234*(5), 5451-5465.

Schlesselman, J. J. (1982). Case-control studies: design, conduct, analysis (Vol. 2). Oxford university press.

Sebastian-delaCruz, M., Gonzalez-Moro, I., Olazagoitia-Garmendia, A., Castellanos-Rubio, A., & Santin, I. (2021). The role of lncRNAs in gene expression regulation through mRNA stabilization. *Non-coding RNA*, *7*(1), 3.

Shatoff, E., & Bundschuh, R. (2020). Single nucleotide polymorphisms affect RNA-protein interactions at a distance through modulation of RNA secondary structures. *PLoS computational biology*, *16*(5), e1007852.

Simantov, K., Goyal, M., & Dzikowski, R. (2022). Emerging biology of noncoding RNAs in malaria parasites. *PLoS Pathogens*, *18*(7), e1010600.

Sinclair, D., Zani, B., Donegan, S., Olliaro, P., & Garner, P. (2009). Artemisinin-based combination therapy for treating uncomplicated malaria. *Cochrane Database of Systematic Reviews*(3).

Smith, J. D. (2014). The role of PfEMP1 adhesion domain classification in Plasmodium falciparum pathogenesis research. *Molecular and biochemical parasitology*, *195*(2), 82-87.

Smith, M. L., & Styczynski, M. P. (2018). Systems biology-based investigation of host–plasmodium interactions. *Trends in parasitology*, *34*(7), 617-632.

Soe, T. N., Wu, Y., Tun, M. W., Xu, X., Hu, Y., Ruan, Y., Win, A. Y. N., Nyunt, M. H., Mon, N. C. N., & Han, K. T. (2017). Genetic diversity of Plasmodium falciparum populations in southeast and western Myanmar. *Parasites & vectors*, *10*, 1-6.

Soulard, V., Bosson-Vanga, H., Lorthiois, A., Roucher, C., Franetich, J.-F., Zanghi, G., Bordessoulles, M., Tefit, M., Thellier, M., & Morosan, S. (2015). Plasmodium falciparum full life cycle and Plasmodium ovale liver stages in humanized mice. *Nature communications*, *6*(1), 7690.

Tadesse, F. G., Lanke, K., Nebie, I., Schildkraut, J. A., Gonçalves, B. P., Tiono, A. B., Sauerwein, R., Drakeley, C., Bousema, T., & Rijpma, S. R. (2017). Molecular markers for sensitive detection of Plasmodium falciparum asexual stage parasites and their application in a malaria clinical trial. *The American journal of tropical medicine and hygiene*, *97*(1), 188.

Taylor-Robinson, D., Jones, K., Garner, P., & Sinclair, D. (2008). Malaria: uncomplicated, caused by Plasmodium falciparum. *BMJ clinical evidence*, *2008*.

Tran, T. M., Jones, M. B., Ongoiba, A., Bijker, E. M., Schats, R., Venepally, P., Skinner, J., Doumbo, S., Quinten, E., & Visser, L. G. (2016). Transcriptomic evidence for modulation of host inflammatory responses during febrile

Plasmodium falciparum malaria. *Scientific Reports*, *6*(1), 31291.

Vallejos-Vidal, E., Reyes-Cerpa, S., Rivas-Pardo, J. A., Maisey, K., Yáñez, J. M., Valenzuela, H., Cea, P. A., Castro-Fernandez, V., Tort, L., & Sandino, A. M. (2020). Single-nucleotide polymorphisms (SNP) mining and their effect on the tridimensional protein structure prediction in a set of immunity-related expressed sequence tags (EST) in Atlantic salmon (Salmo salar). *Frontiers in genetics*, *10*, 1406.

Venkatesan, P. (2024). The 2023 WHO World malaria report. The Lancet Microbe, 5(3), e214.

Wanowska, E., Kubiak, M. R., Rosikiewicz, W., Makałowska, I., & Szcześniak, M. W. (2018). Natural antisense transcripts in diseases: From modes of action to targeted therapies. *Wiley Interdisciplinary Reviews: RNA*, *9*(2), e1461.

WHO. (2000). Severe falciparum malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *94*, 1-90.

WHO. (2015). *Global technical strategy for malaria 2016-2030*. World Health Organization.

WHO. (2019). World malaria report 2019. In *World malaria report 2019*.

WHO. (2020). World malaria report 2020: 20 years of global progress and challenges. *World malaria report 2020: 20 years of global progress and challenges*, 299.

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

Witmer, K., Fraschka, S. A., Vlachou, D., Bartfai, R., & Christophides, G. K. (2020). An epigenetic map of malaria parasite development from host to vector. *Scientific Reports*, *10*(1), 6354.

Wucher, V., Legeai, F., Hédan, B., Rizk, G., Lagoutte, L., Leeb, T., Jagannathan, V., Cadieu, E., David, A., & Lohi, H. (2017). FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic acids research*, *45*(8), e57-e57.

Xia, L., Wu, J., Pattaradilokrat, S., Tumas, K., He, X., Peng, Y.-c., Huang, R., Myers, T. G., Long, C. A., & Wang, R. (2018). Detection of host pathways universally inhibited after Plasmodium yoelii infection for immune intervention. *Scientific Reports*, *8*(1), 15280.

Yan, H., & Bu, P. (2021). Non-coding RNA in cancer. *Essays in biochemistry*, *65*(4), 625-639.

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*, *16*(5), 284-287.

Zhang, T., Xu, X., Jiang, J., Yu, C., Tian, C., & Li, W. (2018). Surveillance of antimalarial resistance molecular markers in imported Plasmodium falciparum

malaria cases in Anhui, China, 2012–2016. *The American journal of tropical medicine and hygiene*, *98*(4), 1132.

Zhao, W., An, Y., Liang, Y., & Xie, X. (2014). Role of HOTAIR long noncoding RNA in metastatic progression of lung cancer. *Eur Rev Med Pharmacol Sci*, *18*(13), 1930-1936.

Zhu, R., Howorka, S., Pröll, J., Kienberger, F., Preiner, J., Hesse, J., Ebner, A., Pastushenko, V. P., Gruber, H. J., & Hinterdorfer, P. (2010). Nanomechanical recognition measurements of individual DNA molecules reveal epigenetic methylation patterns. *Nature nanotechnology*, *5*(11), 788-791.

Zinder, D., Rorick, M. M., Tiedje, K. E., Ruybal-Pesántez, S., Day, K. P., & Pascual, M. (2017). Conservation of single amino-acid polymorphisms in Plasmodium falciparum erythrocyte membrane protein 1 and association with severe pathophysiology. *bioRxiv*, 128983.

APPENDICES

Appendix 1: Ethical Approvals

MCMR-UWZ-C 4 January 2018

MEMORANDUM FOR Director, Human Subjects Protection Branch (HSPB), Walter Reed Army Institute of Research (WRAIR), 503 Robert Grant Ave., Silver Spring, MD 20910-7500

SUBJECT: Approval of the Minimal Risk Human Subjects Research Protocol **WRAIR #2454**

1. I approve the protocol, **WRAIR #2454**, entitled "Epidemiology of Malaria and Drug Sensitivity Patterns in Kenya," (Protocol Version 1.6, dated 21 December 2017), submitted by Ben Andagalu, MD, MSc, Deputy Chief, Malaria Drug Resistance Laboratory, Kenya Medical Research Institute (KEMRI)/Walter Reed Project (WRP).

2. This study is being funded by the U.S. Department of Defense Global Emerging Infections Surveillance and Response System (GEIS), Armed Forces Health Surveillance.

3. This protocol will replace the currently approved protocol WRAIR# 1384, entitled "Epidemiology of Malaria and Drug Sensitivity Patterns in Kenya." This is a combined clinical-laboratory based study to determine *Plasmodium* parasites drug resistance patterns and parasite genetic characteristics over time. The general objectives of this study are as follows: (1) To establish the *ex vivo/in vitro* drug resistance patterns of malaria parasites collected at various study sites in Kenya and to correlate these outcomes with the historical data obtained by the MDR laboratory. Information will be provided to GEIS and the Kenya National Malaria Control Program (NMCP), which may correlate our findings with their *in vivo* results. (2) To correlate genome-wide polymorphisms with antimalarial drug sensitivity tests determined by *in vivo* and *in vitro* methods respectively.

Up to 100 subjects per week will be enrolled from the surveillance sites located at the Kenya Ministry of Health facilities. Approximately 2500 subjects/year will be enrolled from the Kenya Defense Forces site. Subjects 6 months and older will be enrolled in this study.

4. The WRAIR Scientific Review Committee approved this protocol (Version 1.4, dated 14 August 2017) on 6 September 2017. There have been no scientific changes made between protocol Version 1.4, dated 14 August 2017 and Version 1.6, dated 21 December 2017.

5. The protocol (Version 1.5, dated 28 September 2017) was reviewed by the fully convened WRAIR Institutional Review Board (IRB) on 13 December 2017 and approved with stipulations. The protocol was determined to be a minimal risk study as per 32 CFR 219.111 and 45 CFR 46.111. This study also meets the criteria under 45 CFR 46.204, as it is research involving pregnant women, and 45 CFR 46.404, as it involves children participating on a not greater than minimal risk study.

The study team submitted the response and an updated protocol (Version 1.6, dated 21 December 2017) to the WRAIR IRB stipulations on 3 January 2018. The response to the stipulations was determined to be satisfactory on 4 January 2017.

6. The KEMRI Scientific and Ethics Review Unit (SERU) approval is still pending.

The Wellcome Trust Sanger Institute is not engaged in human subjects research as per the memorandum dated 9 October 2017.

The Medical Research and Materiel Command (MRMC) IRB approval/deferral is pending.

The Naval Medical Research Unit – Peru (NAMRU-6) IRB approval/deferral/determination and the Pasteur of Cambodia IRB approval/determination are pending.

7. In accordance with the U.S. Army Medical Research and Materiel Command (USAMRMC) policy 2013-75, this protocol does require review by the USAMRMC, Office of Research Protections (ORP), Human Research Protections Office (HRPO) as this is research conducted by personnel assigned or attached to an established USAMRMC overseas research institution where the research will enroll minors and pregnant women. The USAMRMC ORP HRPO approval will need to be received prior to issuance of the Commander Approval Authorization.

8. The following supporting documentation must be submitted to the WRAIR HSPB prior to the issuance of the WRAIR Commander Approval Authorization:

a. KEMRI SERU approval;
b. USAMRMC ORP HRPO approval;
c. MRMC IRB approval/deferral;
d. NAMRU-6 IRB approval/deferral/determination;
e. Pasteur of Cambodia IRB approval/determination;
f. Translated Informed Consent and Assent Documents into Kiswahili, Luo, Kisii, and Kipsigis; and
g. Translation Verification Certificates for the Informed Consent and Assent Documents translated into Kiswahili, Luo, Kisii, and Kipsigis.

9. The following documents are included as part of this approval:

a. Protocol Version 1.6, dated 21 December 2017;
b. Adult Informed Consent Form (English), Version 1.4, dated 21 December 2017;
c. Assent Form, Ages 13 through 17 Years, Version 1.3, dated 14 August 2017;
d. Data Collection Tool #1, Version 1.0, dated 9 March 2017; and
e. Data Collection Tool #2, Version 1.1, dated 14 August 2017.

10. The WRAIR expiration date will be **13 December 2018**. The PI is responsible for submitting a continuing review report to the KEMRI SERU and the WRAIR IRB in time for the report to be reviewed and accepted/approved prior to the respective expiration dates to avoid an interruption in work. A closeout report will be due to the WRAIR HSPB on **13 December 2027**. No changes, amendments, or addenda may be made to the protocol without prior review and approval by the KEMRI SERU and the WRAIR IRB, as well as, USAMRMC ORP HRPO, as applicable.

11. The point of contact for this action is Teresa R. Soderberg, M.A., RAC, at (301) 319-9438 or Teresa.R.Soderberg.civ@mail.mil.

LISA M. LEE, PHD
Chair, Institutional Review Board
Walter Reed Army Institute of Research

CF:
Ben Andagalu, MD, MSc
Lucas Otieno, MBChB, MMed
Douglas Shaffer, M.D.
Victor Melendez, LTC, MS
Stacey Gondi
MCMR-RP

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel: (254) (020) 2722541, 2713349, 0722-205901, 0733-400003, Fax: (254) (020) 2720030
E-mail: director@kemri.org, info@kemri.org, Website. www.kemri.org

KEMRI/RES/7/3/1 **April 25, 2018**

TO: **DR. BEN ANDAGALU,**
PRINCIPAL INVESTIGATOR

THROUGH: **THE DIRECTOR, CCR,**
NAIROBI 02/05/2018

Dear Sir,

RE: SERU PROTOCOL NO. 3628 *(RESUBMISSION OF INITIAL SUBMISSION):* EPIDEMIOLOGY OF MALARIA AND DRUG SENSITIVITY PATTERNS IN KENYA

Reference is made to your letter dated March 28, 2018 The KEMRI Scientific and Ethics Review Unit (SERU) acknowledges receipt of the revised application on April 6, 2018.

The SERU Secretariat acknowledges receipt of the following documents:

a. WR2454_MDR_protocol_v1.7 dated 12th March 2018, clean and tracked
b. WR2454_MDR_AdultParent_consent v1.5 MoH_12Mar2018 clean and tracked
c. WR2454_MDR_AdultParent_consent_KDF_v 1.0 dated_12Mar2018 clean
d. Removal letter of Maj Jacob Johnson
e. Appendix 1_attempts to find Maj Jacob Johnson
f. Local trabslations of the ICFs, Luo, Kiswahili, Kisii, Kipsigis
g. Certificates of translations for the ICFs.

This is to inform you that the Committee notes that the issues raised during the 272nd Committee B meeting of the SERU held on **February 21, 2018** have been adequately addressed.

Consequently, the study is granted approval for implementation effective this day, **April 25, 2018** for a period of one year. Please note that authorization to conduct this study will automatically expire on **April 24, 2019.** If you plan to continue data collection or analysis beyond this date, please submit an application for continuation approval to SERU by **March 13, 2019.**

You are required to submit any proposed changes to this study to the SERU for review and the changes should not be initiated until written approval from the SERU is received. Please note that any unanticipated problems resulting from the implementation of this study should be brought to the attention of the SERU and you should advise the SERU when the study is completed or discontinued.

You may embark on the study.

Yours faithfully,

FOR: **THE HEAD,**
KEMRI SCIENTIFIC AND ETHICS REVIEW UNIT

In Search of Better Health

DEPARTMENT OF THE ARMY
WALTER REED ARMY INSTITUTE OF RESEARCH
503 ROBERT GRANT AVENUE
SILVER SPRING, MD 20910-7500

REPLY TO ATTENTION OF

FCMR-UWS-HP

MEMORANDUM FOR Director, Human Subjects Protection Branch (HSPB), Walter Reed Army Institute of Research (WRAIR), 503 Robert Grant Avenue, Silver Spring, Maryland 20910-7500

SUBJECT: Continuing Review Report Approval for the Greater than Minimal Risk Human Subjects Research Protocol, **WRAIR #2739**, OHRO Log #E04203.a-c

1. The corrected continuing review report, dated 20 April 2023, for the protocol **WRAIR #2739**, OHRO Log #E04203.a-c, titled, "Clinical Investigation Study to Evaluate the Consistency and Reproducibility of Two Consecutive Mosquito Feeding Assays in Adults with Varying *Plasmodium falciparum* Gametocyte Densities," (Protocol Version 1.4, dated 31 October 2022) and supporting documents, submitted by Hoseah Akala, PhD, Kenya Medical Research Institute (KEMRI)/Walter Reed Project (WRP), Kenya is approved.

2. The continuing review report covers the reporting period from 30 March 2022 through 29 March 2023. This study is closed to enrollment, remains open for the analysis of specimens and data and is awaiting final closure by the Sponsor.

3. As this greater than minimal risk protocol is permanently closed to enrollment and remains open for the analysis of specimens and data, the continuing review report was reviewed and approved via expedited review procedures according to 32 CFR 219.110(b)(1)(i), and 45 CFR 46.110(b)(1)(i), per expedited category 8c, continuing review of research previously approved by the convened IRB where the remaining research activities are limited to data analysis. This study continues to meet the requirements under 32 CFR 219.111 and 45 CFR 46.111.

4. The KEMRI Scientific and Ethics Review Unit (SERU) approved this protocol for continuation until 22 September 2023, per the memorandum dated 23 September 2022.

An Institutional Agreement for Institutional Review Board (IRB) Review (IAIR) between the KEMRI SERU and PATH was established for this study on 25 November 2020, allowing PATH to rely on the KEMRI SERU for ethical review.

To avoid an interruption in work, please provide a copy of the updated ethics board approvals to the WRAIR HSPB as they become available.

5. This review also constitutes the Human Research Protections Official (HRPO) continuing review approval for the participating non-DoD collaborators in accordance with DoDI 3216.02, Section 3.6, (b)5. This HRPO review applies only to this continuing review action; all other reporting and review requirements by the US Army Medical Research and Development Command (USAMRDC), Office of Human and Animal Research Oversight (OHARO), Office of Human Research Oversight (OHRO), remain in effect.

6. This study continues to be sponsored and funded by PATH with support from the Bill and Melinda Gates Foundation.

In Search of Better Health

KENYA MEDICAL RESEARCH INSTITUTE

OFFICE OF THE DIRECTOR RESEARCH & DEVELOPMENT

Tell: +254 020 2722541, 2713349,
0722 205 901, 0733 400 003

P.O. Box 54840-00200, Nairobi
Email: ddrt@kemri.go.ke
Website: www.kemri.go.ke

KEMRI/RES/7/3/1

September 23, 2022

TO: **DR. HOSEA AKALA,**
PRINCIPAL INVESTIGATOR.

THROUGH **THE DEPUTY DIRECTOR, CCR,**
NAIROBI.

Dear Sir,

RE: **KEMRI/SERU/CCR/0173/4082** ***(RESUBMITTED REQUEST FOR ANNUAL RENEWAL)*****: CLINICAL INVESTIGATION STUDY TO EVALUATE THE CONSISTENCY AND REPRODUCIBILITY OF TWO CONSECUTIVE MOSQUITO FEEDING ASSAYS IN ADULTS WITH VARYING *PLASMODIUM FALCIPARUM* GAMETOCYTE DENSITIES.**

Reference is made to your letter dated September 21, 2022. The KEMRI Scientific and Ethics Review Unit (SERU) acknowledges receipt of the revised study documents on September 21, 2022.

The Expedited Review Team notes that the issues it raised on the letter dated **August 04, 2022** have been adequately addressed.

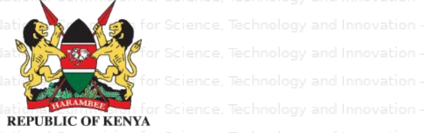
Consequently, the study is **granted approval** for continuation effective **September 23, 2022** through to **September 22, 2023.** Please note that authorization to conduct this study will automatically expire on **September 22, 2023.** If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to SERU by **August 11, 2023.**

You are required to submit any amendments to this protocol and any other information pertinent to human participation in this study to the SERU for review prior to initiation. You may continue with the study.

Yours faithfully,

ENOCK KEBENEI,
THE ACTING HEAD,
KEMRI SCIENTIFIC AND ETHICS REVIEW UNIT

In Search of Better Health



National Commision for Science, Technology and Innovation -

Appendix 2: Characteristics of the filtered reads.

Sample	Unique reads	Duplicate Reads	Sum
ASM0013_R1	346358	210962	557320
ASM0013_R2	330805	226515	557320
ASM0014_R1	134931	297778	432709
ASM0014_R2	101728	330981	432709
ASM0015_R1	486733	257077	743810
ASM0015_R2	427537	316273	743810
ASM0016_R1	414002	126530	540532
ASM0016_R2	327330	213202	540532
ASM0018_R1	864167	147078	1011245
ASM0018_R2	808946	202299	1011245
ASM0020_R1	453098	23709	476807
ASM0020_R2	432847	43960	476807
ASM0025_R1	135895	493504	629399
ASM0025_R2	124007	505392	629399
ASM0028_R1	247110	160652	407762
ASM0028_R2	190824	216938	407762
ASM0031_R1	147345	331553	478898
ASM0031_R2	170967	307931	478898
ASM0033_R1	159473	275509	434982
ASM0033_R2	156645	278337	434982
ASM0034_R1	171463	362398	533861
ASM0034_R2	191696	342165	533861
ASM0036_R1	161989	351612	513601
ASM0036_R2	183550	330051	513601
ASM0037_R1	154683	264854	419537

ASM0037_R2	155043	264494	419537
ASM0038_R1	258801	406650	665451
ASM0038_R2	194109	471342	665451
ASM0039_R1	300495	511881	812376
ASM0039_R2	259823	552553	812376
SM0040_R1	195643	311094	506737
SM0040_R2	181365	325372	506737
SM0041_R1	274668	384009	658677
SM0041_R2	254562	404115	658677
SM0042_R1	246341	405324	651665
SM0042_R2	251716	399949	651665
SM0043_R1	148260	364438	512698
SM0043_R2	161380	351318	512698
SM0045_R1	262023	263505	525528
SM0045_R2	214752	310776	525528
SM0046_R1	227240	420305	647545
SM0046_R2	238729	408816	647545
SM0047_R1	416646	1414273	1830919
SM0047_R2	460052	1370867	1830919
SM0048_R1	477042	228820	705862
SM0048_R2	482948	222914	705862
SM0049_R1	422031	255733	677764
SM0049_R2	392100	285664	677764
SM0050_R1	227806	479277	707083
SM0050_R2	157738	549345	707083

Appendix 3: KEGG pathways associated with the Differentially expressed genes in ASM vs. UM

ID	Description	setSize	enrichmentScore	NES	pvalue	p.adjust	qvalue
pfa03040	Spliceosome	87	-0.4319	-2.86694	0.000102	0.00010202	0.003201573
pfa03008	Ribosome biogenesis in eukaryotes	40	-0.46024	-2.51034	0.000109	0.000108625	0.003201573
pfa03010	Ribosome	85	-0.34146	-2.25827	0.000306	0.00030631	0.00510043
pfa03013	Nucleocytoplasmic transport	25	-0.48127	-2.25286	0.000346	0.000346101	0.00510043
pfa05144	Malaria	39	0.542462	4.352242	0.001221	0.001221001	0.014394962
pfa03050	Proteasome	35	-0.38023	-1.98909	0.002319	0.002318905	0.022782221
pfa01230	Biosynthesis of amino acids	23	-0.43662	-1.98745	0.002801	0.002800794	0.022901849
pfa00071	Fatty acid degradation	9	-0.6016	-1.94072	0.003108	0.003108108	0.022901849
pfa01212	Fatty acid metabolism	15	-0.5056	-1.98092	0.003606	0.00360562	0.023615755
pfa03020	RNA polymerase	23	-0.41421	-1.88544	0.006769	0.006768584	0.036352944
pfa00250	Alanine, aspartate and glutamate metabolism	10	-0.5529	-1.859	0.006784	0.006783719	0.036352944
pfa04146	Peroxisome	13	-0.49154	-1.8306	0.008779	0.008778626	0.043123075
pfa00061	Fatty acid biosynthesis	11	-0.5214	-1.81678	0.010486	0.010486302	0.047549225

pfa01200	Carbon metabolism	33	-0.34701	-1.78349	0.01210 4	0.01210438 6	0.05096583 8
pfa04141	Protein processing in endoplasmic re- ticulum	41	-0.31726	-1.74241	0.01365 1	0.01365113 8	0.05364657 6
pfa03030	DNA replication	6	0.545219	1.67435 8	0.03821 9	0.03821867 7	0.14030944 2
pfa00785	Lipoic acid metabolism	5	0.596565	1.66654 5	0.04046 4	0.04046424 1	0.14030944 2
pfa03420	Nucleotide excision repair	26	-0.33036	-1.57004	0.04446 2	0.04446222 4	0.14560728 4

Appendix 4: single nucleotide polymorphisms Identified in *P. falciparum* virulent genes

CHROM	Chr 2	Chr 2	Chr 2	Chr 2	Chr 2	Chr 2	Chr 2	Chr 5	Chr 5	Chr 5	Chr 5	Chr 5	Chr 5	Chr8	Chr9	Chr9	Chr9	Chr9	Chr9	Chr9
POS	103 494	103 718	103 950	103 969	104 481	104 982	105 037	695 85	699 01	699 61	699 73	700 45	700 51	1366 783	1420 719	1420 960	1420 972	1421 332	1422 363	1422 371
REF	T	A	G	T	T	G	C	T	G	G	G	A	C	G	T	T	C	T	C	A
ALT	C	T	A	C	C	A	T	C	C	T	A	G	A	T	G	A	T	C	T	C
ASM 181		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ASM 183	T	A	G	C	.	G	T	.	.	.	.	.	.	.	G	T	C	T	.	.
ASM 184	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ASM 184	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

ASM 189	T	A	G	C	T	G	C	.	C	G	A	G	A	T	T	T	C	C	C	A
ASM 227	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ASM 265	.	.	G	T	T	G	C	.	.	.	.	.	.	G	.	.	.	.	.	.
ASM 309	.	.	.	.	.	.	.	C	C	T	A	G	A	G	G	T	C	T	T	A
ERR202 8838	T	T	A	T	C	A	C	C	C	T	A	A	A	T	T	A	C	T	C	A
ERR202 8840	T	T	A	T	C	A	C	T	G	T	A	G	A	T	G	A	T	T	C	C
ERR202 8842	T	A	G	T	T	G	T	T	G	T	A	A	A	T	G	T	T	T	C	C
ERR202 8844	T	A	G	T	T	G	C	T	G	G	A	A	C	T	G	T	T	T	C	A
ERR202 8846	T	A	G	T	T	A	T	C	G	T	A	A	A	T	G	T	T	T	C	C
ERR202 8848	C	A	A	T	C	A	C	C	G	T	A	G	A	T	G	A	T	T	C	C
ERR202 8852	T	T	A	T	C	A	C	C	G	G	A	A	C	T	G	A	C	T	C	A
ERR202 8854	T	A	G	T	T	G	T	C	C	G	G	A	C	T	G	A	T	T	C	C
ERR202 8856	T	A	G	T	T	A	T	C	G	T	A	A	C	T	G	A	T	T	C	C
ERR202 8858	C	A	A	T	C	A	C	T	G	T	A	A	A	T	G	A	C	T	C	C
ERR202 8863	C	A	A	T	C	A	C	T	G	G	A	A	A	T	G	A	T	T	C	C
SM0040	C	A	A	T	C	A	C	T	C	T	A	A	C	T	G	A	C	T	C	C
SM0041	.	.	.	.	T	G	T	.	.	.	.	.	.	G	G	T	C	.	.	.

SM0042	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SM0045	T	T	G	C	T	G	T	T	G	G	G	A	C	G	G	T	C	T	T	A
SM0046	T	T	G	C	T	G	T	C	G	T	A	A	C	G	G	T	C	T	T	A
SM0047	T	A	G	T	T	G	C	C	C	G	G	A	C	G	T	T	C	T	C	A
SM0048	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.
SM0049	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SRR11486382	T	A	G	T	T	G	C	T	G	G	G	A	C	G	G	T	C	T	C	A
SRR11486383	T	A	G	T	T	G	C	T	G	G	G	A	C	G	T	A	C	T	C	A
SRR11486384	T	A	G	T	T	G	C	C	G	G	G	A	C	G	G	T	C	T	C	A
SRR11486385	T	T	G	T	T	G	T	C	G	T	A	A	A	G	G	T	C	C	T	A
SRR11486386	T	A	G	T	T	G	C	C	G	G	G	A	C	G	T	A	C	T	C	A
SRR11486387	T	A	G	T	T	G	C	C	G	G	G	A	C	G	T	T	C	T	C	A
SRR11486388	T	A	G	T	T	G	C	C	.	G	G	G	C	G	T	A	C	T	C	A
SRR11486389	T	A	G	T	T	G	C	C	G	G	G	A	C	G	T	T	C	T	C	A
SRR11486391	T	A	G	T	T	G	C	T	G	G	G	A	C	G	T	T	C	T	C	A
SRR11486392	T	A	G	T	T	G	C	C	G	G	G	A	C	G	G	T	C	T	C	A
SRR11486393	T	A	G	T	T	G	C	C	G	G	G	A	C	G	G	T	C	T	C	A
SRR11486394	T	A	G	T	T	G	C	C	G	T	A	G	A	G	T	A	C	T	C	A
SRR11486395	T	A	G	T	T	G	C	C	G	G	G	A	C	G	G	T	C	T	C	A

SRR11486396	T	T	G	T	T	G	T	C	G	G	G	A	C	G	G	T	C	T	C	A
SRR11486397	T	A	G	T	T	G	C	C	G	G	G	G	C	G	G	T	C	T	C	A
SRR11486398	T	A	G	T	T	G	C	C	G	G	A	G	A	G	T	A	C	T	C	A
SRR11486399	T	A	G	T	T	G	C	T	G	G	G	G	C	G	T	T	C	T	C	A
SRR11486400	T	A	G	T	T	G	C	T	G	G	G	G	C	G	T	T	C	T	C	A
SRR11486401	T	A	G	T	T	G	C	C	G	G	G	A	C	G	T	T	C	T	C	A
SRR11486402	T	A	G	T	T	G	C	C	G	G	G	A	C	G	G	T	C	C	C	A
SRR11486403	T	A	G	T	T	G	C	C	G	G	G	A	C	G	T	T	C	T	C	A
SRR11486404	T	A	G	T	T	G	C	T	G	G	G	A	C	G	T	T	C	T	C	A
SRR11486405	T	T	G	T	T	G	T	T	G	G	G	A	A	G	G	T	C	C	C	A