UNIVERSITY OF RWANDA

MOLECULAR CLUSTERING OF CLINICAL MALARIA INFECTIONS BASED ON AMA1 GENE IN RWANDA: IMPLICATIONS ON ARTEMISININ RESISTANCE AND POTENTIAL VACCINE DEVELOPMENT

2025

NDACYAYISENGA Jean Claude

MOLECULAR CLUSTERING OF CLINICAL MALARIA INFECTIONS BASED ON AMA1 GENE IN RWANDA: IMPLICATIONS ON ARTEMISININ RESISTANCE AND POTENTIAL VACCINE DEVELOPMENT

By

NDACYAYISENGA Jean Claude: 222019985

Dissertation submitted in fulfilment of the requirements for the degree :

MASTER OF SCIENCE IN BIOTECHNOLOGY

In the department of Biology, School of Science

College of Science and Technology

at

The University of Rwanda

Supervisor: Dr. Omar MUKAMA

Co-supervisors: Dr. Jean Pierre MUSABYIMANA

Dr. Jean Paul NSHIZIRUNGU

Kigali, Rwanda 2025

DECLARATION

I, NDACYAYISENGA Jean Claude, hereby declare that this research project submitted to the University of Rwanda for the degree Master of Science in Biotechnology is my own original work and has not been submitted before to any institution by myself or any other person in fulfilment of the requirements to the award of any degree or any other qualification.

NDACYAYISENGA Jean Claude

Signature

DEDICATION

I dedicate this dissertation to:

The University of Rwanda, College of Science and Technology, School of Science, Department of Biology.

The European Union and ENABEL, the Belgian development cooperation agency, whose support through the KWIGIRA project funded the Master of Science in Biotechnology program at the University of Rwanda.

My supervisors, Dr. Omar MUKAMA, Dr. Peter MUSABYIMANA, and Dr. Jean Paul NSHIZIRUNGU, for their guidance and mentorship.

Professor Antoine NSABIMANA, Coordinator of the Master of Science in Biotechnology program at the University of Rwanda.

The professors who contributed to the teaching of the Master of Science in Biotechnology program at the University of Rwanda. Their dedication and expertise enriched the program, with contributions from the Université Libre de Bruxelles (ULB), the Université Catholique de Louvain, and the Université de Liège (Belgium); the Université de Grenoble Alpes (France); the University of Gothenburg (Sweden); New York University, Abu Dhabi campus (United Arab Emirates); and the University of Rwanda.

The professors and entire research team at the International Centre for Genetic Engineering and Biotechnology (ICGEB), whose collaboration and support shaped the research experience.

I remain profoundly grateful for the invaluable support that made this academic journey possible. May Almighty God abundantly bless all who contributed to my pursuit of a master's degree.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Dr. Omar MUKAMA, and co-supervisors, Dr. Jean Paul NSHIZIRUNGU and Dr. Jean Pierre MUSABYIMANA, for their invaluable guidance, expertise, and steadfast support throughout this academic journey. Their insightful feedback, constructive criticism, and continuous encouragement have significantly shaped the direction and quality of this research.

My sincere appreciation also goes to Jean Paul HATEGEKIMANA and RUGAMBA Alexis for their collaboration, thoughtful discussions, and meaningful contributions, which greatly enriched the scope and depth of this work.

I am profoundly grateful to the ARMEA Project, the Enabel EU Project Kwigira, and the BK Foundation, as well as to the University of Rwanda – School of Science and Technology, for providing the resources, scholarship support, and a nurturing environment that made this research possible. Their commitment to advancing scientific inquiry, innovation, and capacity building in Rwanda is deeply appreciated.

Finally, I extend heartfelt thanks to all individuals who supported this work, whether directly or indirectly, and whose names may not be explicitly mentioned here. Your kindness, encouragement, and assistance have been truly invaluable. This dissertation is the product of collective effort, and I remain sincerely grateful to everyone who contributed to its realization.

List of figures

Figure 1: shows the multiple sequence alignment of the Plasmodium falciparum AMA1 gene across 29 barcoded clinical isolates from Rwanda. Conserved nucleotide sites are represented in green, while polymorphic loci appear in variant colors (e.g., purple, yellow), indicating nucleotide substitutions at specific genomic positions. 22

Figure 2: presents the multiple sequence alignment of the PfK13 gene across the same 29 P. falciparum isolates. Highly conserved regions are displayed in green, while variable loci, especially within the propeller domain, are highlighted in other colors. 23

Figure 3: This bar chart presents the frequency distribution of missense mutations in the PfAMA1 gene across 29 Plasmodium falciparum isolates. Each bar represents the percentage of isolates with a specific amino acid substitution, separated by district of origin Huye (blue) and Kirehe (orange). Frequencies are calculated as a percentage of the total sample size (n=29). 27

Figure 4: This grouped bar chart compares the total number of missense mutations identified in the PfAMA1 and PfK13 genes across clinical isolates from Huye and Kirehe districts. The height of each bar represents the count of Figure unique mutations detected in each gene per district, based on 29 samples. 29

Figure 5: Distribution of AMA1 and K13 Missense Mutations in P. falciparum Clinical Isolates from Huye and Kirehe Districts, Rwanda 30

Figure 6: Linkage Disequilibrium Pattern Across the AMA1 Gene (r^2 Heatmap). 31

Figure 7: Linkage Disequilibrium Pattern Across the AMA1 Gene (r^2 Heatmap). 32

Figure 8: Sliding Window Analysis of Tajima’s D across the *AMA1* Gene in Plasmodium falciparum 35

Figure 9: Phylogenetic trees of Plasmodium falciparum clinical isolates based on the PfAMA1 (top) and PfK13 (bottom) gene sequences. 37

List of tables

Table 1: (PfK13):Mutations in the PfK13 gene from 29 P. falciparum samples in Huye and Kirehe. R561H, a proven resistance marker, appeared only in Kirehe (10.3%, mixed form). G449A (6.9%, fixed form) was found only in Huye, and N490T (3.4%, mixed form) only in Kirehe. These patterns show possible local spread of resistance in Kirehe and unique low-frequency changes in each district. 24

Table 2: (PfAMA1):Mutations in the PfAMA1 gene from Huye and Kirehe. A total of 24 amino acid changes were found, most rare (3.4%), with some recurring—E187N (10.3%), K206E (6.9%), and Y175D (6.9%). Most were in mixed form and showed a mix of shared and district-specific patterns, indicating immune-driven diversity and regional parasite differences. 26

Table 3: Summary of Genetic Diversity and Neutrality Tests for AMA1 and PfK13 Genes 33

List of Acronyms and Abbreviations

ACTs: Artemisinin-based Combination Therapies

AE: Elution Buffer (Qiagen QIAamp DNA Mini Kit)

AL: Artemether-Lumefantrine

AMA1: Apical Membrane Antigen 1

AM: Artesunate-Mefloquine

AP: Artesunate-Pyronaridine

AS: Artesunate

BCF: Binary Variant Call Format

BWA: Burrows-Wheeler Aligner

CHW: Community Health Worker

CLSI: Clinical and Laboratory Standards Institute

DAPC: Discriminant Analysis of Principal Components

DHA-PPQ: Dihydroartemisinin-Piperaquine

DNA: Deoxyribonucleic Acid

DRC: Democratic Republic of Congo

EDTA: Ethylenediaminetetraacetic Acid

GC Content: Guanine-Cytosine Content

HRP2: Histidine-Rich Protein 2

IRS: Indoor Residual Spraying

LD: Linkage Disequilibrium

LLINs: Long-Lasting Insecticidal Nets

MSA: Multiple Sequence Alignment

MVP: Malaria Vaccine Program

NRL: National Reference Laboratory

ONT: Oxford Nano pore Technologies

PCA: Principal Component Analysis

PCR: Polymerase Chain Reaction

Pf: Plasmodium falciparum

PfK13: Plasmodium falciparum Kelch 13 Gene

RBC: Rwanda Biomedical Centre

RDT: Rapid Diagnostic Test

SNP: Single Nucleotide Polymorphism

sWGA: Selective Whole Genome Amplification

WHO: World Health Organization

θ (Theta): Watterson's Estimator of Nucleotide Diversity

ABSTRACT

Background: Malaria caused by *Plasmodium falciparum* remains a major challenge in Rwanda, where genetic variability and emerging resistance threaten control strategies. This study examined molecular clustering of clinical *P. falciparum* infections based on the AMA1 gene and explored their association with PfK13 mutations and geographic distribution.

Methods: A cross-sectional study analyzed 29 *P. falciparum*-positive isolates from Huye and Kirehe districts. Genomic DNA was sequenced targeting AMA1 and PfK13. Sequence data underwent quality control, SNP calling, and annotation. Genetic clustering was assessed using phylogenetic trees, haplotype networks, and principal component analysis (PCA). Neutrality tests (Tajima's D, Fu & Li's D) evaluated evolutionary pressures. Associations between AMA1 clusters and PfK13 mutations were investigated across regions.

Results: AMA1 showed high polymorphism with 2,494 segregating sites and a Tajima's D of +1.158, consistent with balancing selection. PCA and phylogenetic analyses revealed distinct haplotype clusters, indicating sub-structuring. PfK13 displayed limited variation (Tajima's D = –0.071), reflecting near-neutral evolution. The R561H mutation, linked to artemisinin resistance, appeared only in Kirehe and co-occurred with specific AMA1 haplotypes, suggesting localized clustering of resistant lineages.

Conclusion: Findings demonstrate molecular clustering of *P. falciparum* based on AMA1 diversity, with important implications for vaccine development and resistance monitoring. Integrating molecular and geographic data can enhance early detection of resistance hotspots and guide targeted interventions.

Keywords: *Plasmodium falciparum*, AMA1, PfK13, artemisinin resistance, molecular clustering, Rwanda, haplotype diversity, vaccine development

Table of Contents

DECLARATION ... iii

DEDICATION ... iv

ACKNOWLEDGEMENTS ... v

List of Acronyms and Abbreviations ... viii

ABSTRACT ... x

CHAPTER 1: INTRODUCTION ... 1

1.1. Background ... 1

1.2. Problem statement ... 3

1.3. Conceptual Framework Linking AMA1 Genetic Diversity, Artemisinin Resistance, and Vaccine Development Potential in Plasmodium falciparum from Rwanda ... 4

1.4. Justification of the study ... 5

1.5. Study objectives ... 6

1.5.1. General objective ... 6

1.5.2. Specific objectives ... 6

1.6. Research questions ... 6

1.7. Research hypotheses ... 6

1.7.1. General Hypothesis ... 6

1.7.2. Specific Hypotheses ... 7

1.8. Organization of the study ... 7

CHAPTER 2: LITERATURE REVIEW ... 8

2.1. Theoretical Literature Review ... 8

2.2 Conceptual framework ... 10

CHAPTER 3. METHODOLOGY ... 12

3.1. Study Design ... 12

3.2. Study Sites and Population ... 12

3.3. Selection of Participants ... 13

3.3.1. Sample Size Determination ... 13

3.3.2. Participant Recruitment and Sample Processing ... 14

3.3.3. Training of Data Collectors ... 15

3.3.4. Study Variables ... 15

3.4. Sample Collection and Processing 16
3.5. Genomic DNA Extraction 16
3.6. Whole Genome Amplification (WGA) 16
3.7. Molecular Amplification and Sequencing 17
3.7.1. Target Gene Selection 17
3.7.2. Polymerase Chain Reaction (PCR) Amplification 17
3.7.3. Amplicon Verification and Purification 17
3.7.4. DNA Sequencing 17
3.8. Bioinformatics and Genetic Data Analysis 18
3.8.1. Raw Data Preprocessing 18
3.8.2. SNP Calling and Variant Annotation 18
3.8.3. Amino Acid Translation and Domain Mapping 19
3.8.4. Haplotype Construction and Visualization 19
3.8.5. Genetic Diversity and Neutrality Metrics 19
3.8.6. Linkage Disequilibrium (LD) and Principal Component Analysis (PCA) 20
3.8.7. Phylogenetic and Network Analysis 20
3.8.8. Co-occurrence Analysis of PfK13 and PfAMA1 Variants 20
3.9. Data Analysis 20
3.10. Ethical Considerations 21
CHAPTER 4: RESULTS AND DISCUSSION 22
CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS 40
5.1. Summary of Findings 40
5.2. Conclusions 41
5.3. Recommendations 41
5.4. Limitations and Future Work 41
5.5. Closing Statement 42
References 43

CHAPTER 1: INTRODUCTION

1.1.Background

Malaria remains one of the most significant and persistent public health challenges in the world. Despite remarkable progress over the past two decades in reducing the global malaria burden, the disease continues to affect millions, particularly in sub-Saharan Africa. According to the 2024 World Health Organization (WHO) World Malaria Report, malaria caused an estimated 249 million clinical episodes and over 600,000 deaths in 2022 alone, with the vast majority of cases concentrated in low- and middle-income countries.

In Rwanda, malaria transmission varies spatially and temporally, with the highest endemicity observed in the Eastern and Southern provinces. These regions experience prolonged transmission seasons, favorable climatic conditions for mosquito proliferation, and difficulties in maintaining consistent vector control measures. The Rwanda Biomedical Center (RBC) has prioritised malaria control through strategies including mass LLIN distributions, IRS, community case management, and improved access to ACTs.

Artemisinin-based combination therapies (ACTs) have played a critical role in reducing malaria morbidity and mortality since being adopted as first-line treatment for P. falciparum infections. ACTs combine a fast-acting artemisinin derivative with a longer-acting partner drug to ensure complete parasite clearance (Ashley et al., 2014; Uwimana et al., 2021). Due to artemisinin's rapid action but short half-life (1–3 hours), the partner drug is essential to prevent recrudescence and the development of drug resistance.

The emergence and spread of artemisinin resistance, initially reported in the Greater Mekong Subregion, pose a serious challenge to malaria control and elimination efforts worldwide (Ariey et al., 2014).

The main artemisinin-based combination therapies (ACTs) currently in use across Africa include artemether-lumefantrine (AL), artesunate-amodiaquine (AA), dihydroartemisinin-piperaquine (DHA-PPQ), artesunate-mefloquine (AM), and, more recently, artesunate-pyronaridine (AP) (Ashley et al., 2014).

Artemisinin resistance is characterised clinically by delayed parasite clearance and is associated molecularly with mutations in the P. falciparum Kelch 13 (PfK13) propeller domain. Key validated

mutations, including C580Y, R561H, and others, have been implicated in reduced artemisinin susceptibility (Ariey et al., 2014; Uwimana et al., 2021).

While Southeast Asia has been the epicenter of resistance, recent molecular surveillance has documented the emergence of PfK13 mutations in Africa. In Rwanda, the R561H mutation has been detected and associated with delayed parasite clearance following ACT treatment (Uwimana et al., 2021). Similarly, candidate mutations have been observed in Uganda, highlighting a potential eastward spread of resistance across the continent. In contrast, studies from the Democratic Republic of Congo (DRC) report the absence of validated resistance-associated PfK13 mutations, though novel non-synonymous mutations have been identified, warranting further investigation (Yobi et al., 2020).

Besides drug resistance, other biological and epidemiological factors complicate malaria control in Rwanda. These include asymptomatic parasite carriage, cross-border transmission, HRP2 gene deletions that affect rapid diagnostic tests, and limited genomic surveillance infrastructure. Furthermore, the highly polymorphic nature of parasite antigens, such as the Apical Membrane Antigen 1 (AMA1), presents challenges for both vaccine development and transmission tracking.

AMA1 plays a critical role in the erythrocyte invasion process and is a leading blood-stage malaria vaccine candidate. However, its extreme genetic diversity, driven by host immune pressure and recombination, leads to substantial antigenic variation. Studying AMA1 polymorphisms offers a dual benefit: it informs vaccine design by identifying immunodominant epitopes, and it facilitates analysis of parasite population structure through molecular clustering.

This study focuses on the molecular clustering of Plasmodium falciparum infections based on AMA1 haplotypes in Rwanda, while also evaluating the co-occurrence of PfK13 mutations associated with artemisinin resistance. By integrating parasite genotyping, spatial distribution, and resistance profiling, the research aims to contribute to national and regional efforts in malaria control, resistance containment, and informed vaccine strategies.

1.2.Problem statement

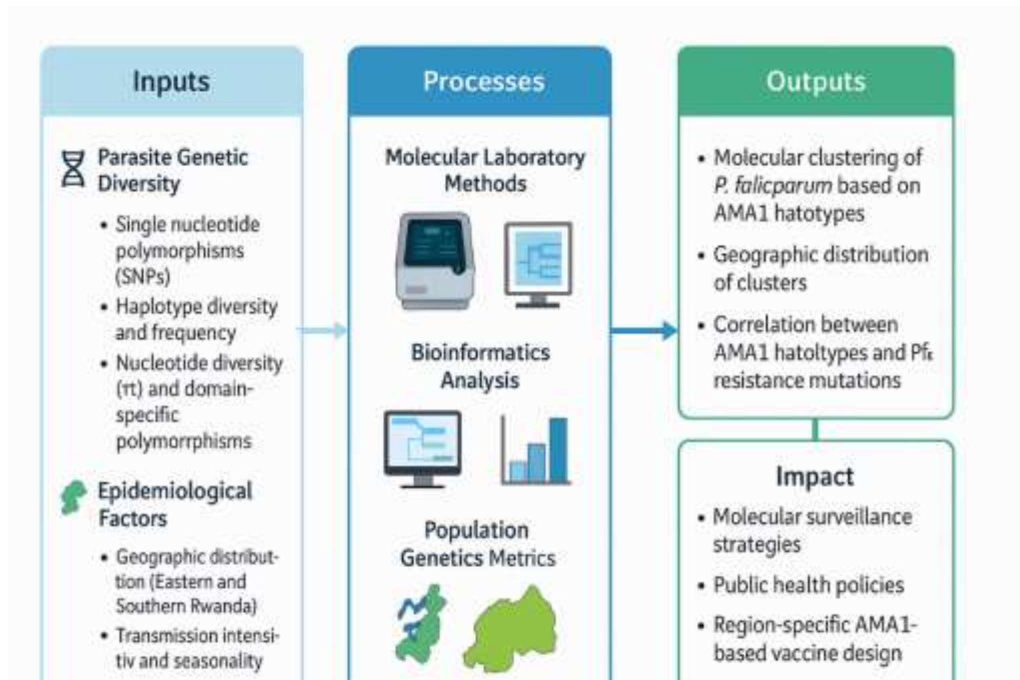
Despite major investments in malaria prevention and treatment, Rwanda continues to experience a high malaria burden, particularly in its Eastern and Southern provinces. These regions exhibit persistent transmission even in the face of widespread deployment of long-lasting insecticidal nets (LLINs), indoor residual spraying (IRS), and ACTs. More concerning is the increasing documentation of artemisinin resistance-associated mutations, notably the R561H mutation, which has been confirmed in Rwanda and linked to reduced sensitivity to artemisinin derivatives.

The absence of validated PfK13 mutations in other countries such as the Democratic Republic of Congo (DRC) raises questions about local transmission dynamics and the possible introduction or independent emergence of resistant strains (Straimer et al., 2022). While some resistance hotspots in East Africa exhibit common genotypes, Rwanda shows signs of unique mutation profiles. Moreover, diagnostic challenges such as HRP2-negative strains further complicate surveillance and case management (Mihreteab et al., 2023; Hodder et al., 2001).

Despite this, few studies have comprehensively examined the genetic diversity of malaria parasites circulating in Rwanda. Specifically, limited data exist on AMA1 polymorphisms, which serve as molecular markers of strain diversity and immune evasion. Given that AMA1 is also a candidate for blood-stage malaria vaccine development, understanding its haplotype structure provides dual benefits: monitoring strain transmission and guiding vaccine formulation.

Therefore, this study aims to fill critical gaps by characterizing AMA1 diversity and clustering patterns and linking them to K13 mutations and geographic origin. By doing so, it can identify transmission networks, assess resistance dissemination, and support policy interventions grounded in local evidence.

1.3.Conceptual Framework Linking AMA1 Genetic Diversity, Artemisinin Resistance, and Vaccine Development Potential in Plasmodium falciparum from Rwanda



The conceptual framework shows how parasite genetic diversity, artemisinin resistance markers, and epidemiological factors work together to influence malaria transmission and control strategies in Rwanda. It links these inputs with molecular laboratory analysis, bioinformatics, and population genetics to generate actionable outputs. The process starts with characterizing AMA1 genetic polymorphisms, haplotype diversity, and PfK13 resistance mutations, alongside geographic and transmission data. These analyses yield molecular clustering patterns, map the geographic distribution of AMA1 haplotypes, and uncover associations with PfK13 mutations. By translating these findings into practical applications, the framework guides evidence-based public health policies, strengthens molecular surveillance systems, and supports the design of region-specific AMA1-based malaria vaccines. In essence, it connects the genetic and epidemiological features of Plasmodium falciparum to analytical workflows, ensuring that complex molecular data lead to targeted, impactful interventions.

1.4.Justification of the study

Molecular surveillance has emerged as a critical pillar in contemporary malaria control strategies, especially in regions where early signals of antimalarial drug resistance are appearing. In Rwanda, the identification of validated *Plasmodium falciparum* kelch 13 (*PfK13*) mutations known markers of artemisinin resistance has heightened the urgency for regionally tailored molecular data to support precision public health interventions. National Malaria Control Programs (NMCPs) increasingly rely on real-time, high-resolution data on parasite genetics and population structure to guide treatment policies, allocate resources efficiently, and monitor the impact of interventions.

This study is both timely and significant for several reasons:

1. Leveraging AMA1 for Population Clustering: It utilizes *Apical Membrane Antigen 1* (AMA1) gene diversity as a molecular tool to cluster *P. falciparum* infections based on haplotype patterns. AMA1 is a highly polymorphic surface protein involved in erythrocyte invasion, and its variability reflects immune selection and transmission dynamics. By characterizing haplotype diversity and clustering patterns, the study will enhance understanding of parasite population structure and gene flow within and across regions.

2. Linking Genetic Clusters to Drug Resistance: By examining the association between AMA1 haplotypes and *PfK13* mutations, the study aims to uncover potential co-evolutionary dynamics between immune evasion mechanisms and emerging drug resistance. Detecting whether specific AMA1 clusters are enriched for *PfK13* mutations can provide early warnings of adaptive parasite lineages with dual survival advantages—immune escape and treatment resistance.

3. Alignment with Global and National Priorities: The 2024 World Health Organization (WHO) Malaria Report underscores the need to expand genomic surveillance in the African region as a key strategy to combat rising resistance and stagnating malaria progress. This study directly addresses that call by generating country-specific evidence on the genetic epidemiology of malaria parasites in Rwanda. Integrating molecular data with geographic origin of samples will enable spatial mapping of transmission and resistance patterns, thus equipping the NMCP with the intelligence to deploy targeted and adaptive control measures.

1.5.Study objectives

1.5.1.General objective

To investigate the molecular clustering patterns of clinical *Plasmodium falciparum* infections based on genetic polymorphisms in the AMA1 gene in Rwanda, and to assess the potential implications of these clusters on artemisinin resistance dynamics and malaria vaccine development.

1.5.2.Specific objectives

To characterise the genetic polymorphisms and haplotype diversity of the AMA1 gene among P. falciparum isolates in Rwanda.

To perform molecular clustering of clinical malaria infections based on AMA1 haplotype structures.

To correlate/evaluate the association between the K13 mutation with AMA1 clusters and geographic regions/study area

1.6.Research questions

What are the genetic polymorphisms and the extent of haplotype diversity in the AMA1 gene among Plasmodium falciparum clinical isolates in Rwanda?

How are P. falciparum infections clustered based on AMA1 haplotype structures, and what do these clusters reveal about parasite population structure and transmission dynamics?

Is there a significant association between PfK13 mutations and specific AMA1 haplotype clusters or geographic regions within Rwanda?

1.7.Research hypotheses

1.7.1.General Hypothesis

There is significant genetic clustering among Plasmodium falciparum clinical isolates in Rwanda, based on the AMA1 gene, which correlates with PfK13 mutations associated with artemisinin resistance and has implications for vaccine development.

1.7.2.Specific Hypotheses

1. H_1: There is significant genetic polymorphism and high haplotype diversity in the AMA1 gene among Plasmodium falciparum isolates in Rwanda.

3. H_2: Clinical P. falciparum infections in Rwanda form distinct molecular clusters based on AMA1 haplotype structures, reflecting population stratification and transmission patterns.

4. H_3: There is a statistically significant association between PfK13 mutations and specific *AMA1* haplotype clusters and/or geographic regions, suggesting possible co-evolution of drug resistance and immune evasion.

1.8.Organization of the study

This study is subdivided into six main chapters; the first chapter is the introduction containing: Background representing the situation of severe malaria burden and its updated statistics, the problem statement which represents the motivation to conduct this study, the justification of the study, research objectives, questions and this one of organization. The second chapter is the literature review which contains concepts relevant to the study and is subdivided into a severe malaria overview, theoretical literature review part and theoretical framework of the factors part. The third chapter shows the methodology that was followed all along with this study. The fourth chapter is the results section which shows the findings of the study. The fifth is the discussion and lastly conclusion and recommendations.

CHAPTER 2: LITERATURE REVIEW

2.1. Theoretical Literature Review

Malaria, primarily caused by *Plasmodium falciparum*, remains a major global health challenge, particularly in sub-Saharan Africa. Rwanda has made significant strides in malaria control, yet the persistence of transmission hotspots and the emergence of artemisinin resistance threaten these gains. The apical membrane antigen 1 (AMA1) is a polymorphic protein essential for merozoite invasion into erythrocytes and is a leading candidate for malaria vaccine development. This literature review contextualizes AMA1 polymorphism and molecular clustering in *P. falciparum*, links it with artemisinin resistance, and outlines the implications for vaccine strategies, with a focus on Rwanda and East Africa.

AMA1 is a key surface protein expressed during the late schizont stage of the *P. falciparum* life cycle. Structurally, it is composed of three ectodomains (DI, DII, DIII) and is a highly polymorphic gene due to strong immune selection pressure. The variation in AMA1 has been documented in multiple studies, demonstrating extensive allelic diversity and positive selection, particularly within domain I (DI), where immune pressure is most intense (Ariey et al., 2014; Ashley et al., 2014).

In a comprehensive study by Kusi et al. (2017), more than 200 AMA1 haplotypes were identified in Ghana, highlighting the challenge of vaccine design due to high polymorphism. Similar findings in Southeast Asia and Papua New Guinea (Barry et al., 2009) have shown that AMA1 diversity contributes to immune evasion and has implications for strain-specific immunity.

Recent studies in Africa, particularly Rwanda, are beginning to characterize AMA1 variation. These studies show localized haplotype distributions, suggesting geographical clustering that may correspond with transmission intensity and immune history of the population (Ndoli et al., 2024).

Haplotype diversity is a key parameter in understanding the genetic structure of malaria parasites. AMA1 haplotypes exhibit high recombination and frequency-dependent selection, making them ideal markers for analyzing population diversity. Molecular clustering based on AMA1 haplotypes can identify transmission patterns, parasite lineages, and potential vaccine escape mutants.

According to Barry et al. (2009), clustering AMA1 sequences into genetic groups can reveal shared ancestry or recent transmission links. Techniques like STRUCTURE, DAPC (Discriminant Analysis of Principal Components), and network-based clustering have been employed to map the genetic landscape of *P. falciparum* populations.

In Rwanda, few studies have comprehensively applied molecular clustering using AMA1; however, the framework has been applied to MSP1 and other loci to demonstrate spatial clustering of genetic variants (Rulisa et al., 2013). Applying similar approaches to AMA1 can reveal micro-epidemiological patterns crucial for targeting interventions.

Artemisinin resistance in *P. falciparum* is largely attributed to mutations in the Kelch13 (K13) propeller domain. Initially identified in Southeast Asia, K13 mutations have now been confirmed in several African countries, including Rwanda (Ariey et al., 2014). Rwanda has emerged as a key site for resistance evolution, with mutations such as R561H, 469F, and A675V found in southern and eastern regions (Uwimana et al., 2021a). These mutations are associated with delayed parasite clearance but have not yet resulted in high treatment failure rates due to the continued efficacy of partner drugs like lumefantrine (Uwimana et al., 2021b).

The linkage between AMA1 clustering and K13 mutations remains underexplored, but theoretically, shared transmission networks or selective pressure environments could result in co-association. Clusters with high AMA1 diversity may also harbor diverse K13 backgrounds, providing insight into how resistance spreads within genetically distinct parasite populations.

Studies conducted in Ruhuha and Huye districts of Rwanda demonstrate spatial clustering of malaria infections. Spatial autocorrelation and SaTScan analyses have revealed persistent hotspots even in areas considered low-endemic (Rulisa et al., 2013). These clusters are often associated with environmental risk factors (e.g., swamps, rice fields), socio-economic status, and intervention coverage.

Integrating spatial clustering of AMA1 haplotypes with geographic K13 mutation mapping can enhance understanding of how resistant and immunologically evasive strains spread across regions. The presence of AMA1 and K13 variants in overlapping hotspots could indicate regions where parasite evolution is most dynamic, providing key targets for surveillance and control.

AMA1 has long been considered a promising vaccine target due to its crucial role in erythrocyte invasion. However, high polymorphism poses a challenge. Most vaccine efforts, such as the FMP2.1/AS02A AMA1-based vaccine, showed limited efficacy due to strain-specific responses (Dutta et al., 2009). Multivalent approaches combining multiple AMA1 alleles have shown promise in increasing immune coverage but at the cost of complexity.

Analyzing local AMA1 haplotypes in Rwanda can inform the design of regionally adapted multivalent vaccines. By identifying dominant haplotypes and their immune relevance, a tailored AMA1-based vaccine could be more effective than globally designed formulations. Additionally, molecular clustering can identify haplotypes under positive selection, which may indicate immunologically dominant strains relevant for vaccine targeting. Studies in Papua New Guinea and Mali (Barry et al., 2009; Ouattara et al., 2020) have demonstrated that some AMA1 clusters are more immunogenic than others.

An integrated approach combining AMA1 haplotype diversity, K13 mutation tracking, and spatial genomics offers a powerful toolset for malaria control and vaccine development. This aligns with WHO's emphasis on genomic surveillance to inform real-time intervention planning (World Health Organization, 2023b). In Rwanda, where both vaccine-preventable variants and resistance mutations are emerging, such integration is critical. The ability to molecularly cluster infections based on AMA1 while simultaneously evaluating K13 mutation profiles and geography can provide a nuanced picture of malaria evolution and control dynamics.

2.2 Conceptual framework

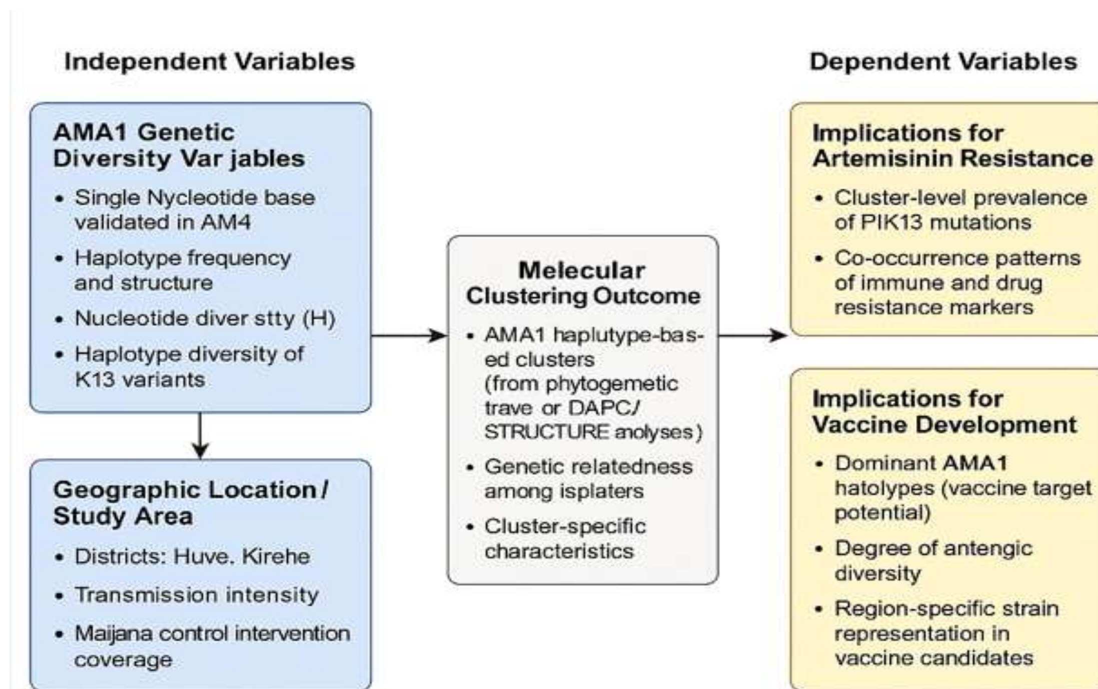
The conceptual framework demonstrates how genetic diversity and geographic factors interact to influence molecular clustering outcomes in *Plasmodium falciparum*, with important implications for artemisinin resistance surveillance and malaria vaccine development. It organizes the study variables into three interconnected components: independent variables, molecular clustering outcomes, and dependent variables.

On the left, the independent variables are categorized into genetic and geographic factors. Genetic diversity metrics include validated single nucleotide polymorphisms (SNPs) within the *PfAMA1* gene, measures of haplotype frequency, nucleotide diversity (π or H), and the haplotype diversity of *PfK13* variants, which are relevant markers for antimalarial resistance. These genetic data form the foundation for identifying population structure and evolutionary dynamics. Complementing this, geographic variables such as the study districts (Huye and Kirehe), levels of malaria transmission, and intervention coverage (e.g., use of insecticide-treated nets and indoor residual spraying) provide essential epidemiological context for interpreting the genetic findings.

At the core of the framework lies the molecular clustering outcome. Using computational methods like phylogenetic analysis, Discriminant Analysis of Principal Components (DAPC), or Bayesian clustering

tools (e.g., STRUCTURE), isolates are grouped into AMA1 haplotype-based clusters. These clusters reflect genetic relatedness, population differentiation, and potential region-specific lineage expansion or immune selection patterns.

On the right, the framework outlines the dependent variables, which focus on two key applications. First, in relation to artemisinin resistance, molecular clusters can be analyzed for the distribution and prevalence of *PfK13* mutations. This enables the identification of resistant subpopulations and the assessment of co-occurrence patterns between immune and drug resistance markers. Second, for vaccine development, understanding which AMA1 haplotypes dominate within clusters can help prioritize antigen variants for inclusion in multivalent vaccine formulations. Additionally, the extent of genetic diversity and the identification of region-specific strains support the tailoring of vaccines to local transmission settings.



CHAPTER 3. METHODOLOGY

3.1. Study Design

This study employed a cross-sectional molecular epidemiological design to investigate the genetic diversity and population structure of Plasmodium falciparum clinical isolates collected from malaria-endemic districts in Rwanda. The focus was on the apical membrane antigen-1 (PfAMA1) gene and the Kelch 13 (PfK13) propeller domain, which are relevant for vaccine design and artemisinin resistance surveillance, respectively. This design facilitated concurrent collection and molecular analysis of parasite genomic data across geographically distinct populations to infer selection pressures, mutation prevalence, transmission dynamics, and resistance patterns.

3.2. Study Sites and Population

The research was conducted in Huye District (Southern Province) and Kirehe District (Eastern Province). These districts were selected based on documented differences in malaria transmission intensity and accessibility to clinical populations.

Participants were recruited from health centres and community-level services using the following:

Inclusion criteria:

Confirmed P. falciparum infection via RDT and/or microscopy

Aged ≥6 months

Residency in the district for ≥6 months

Willingness to provide informed consent

Exclusion criteria included:

Mixed-species malaria infections

Severe malaria requiring inpatient treatment

Refusal to consent

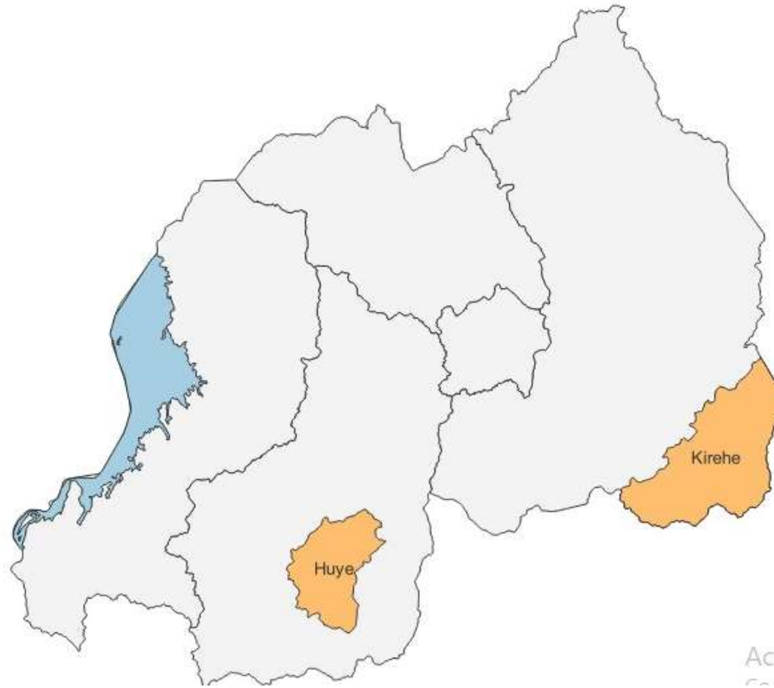


Fig.1 Map of Rwanda showing the selected districts

3.3. Selection of Participants

3.3.1. Sample Size Determination

The sample size for this study was calculated based on the estimated prevalence of *PfK13* gene mutations among malaria-positive individuals in endemic regions of Rwanda. Assuming an expected prevalence of 30% [5], which maximizes the required sample size and ensures sufficient power, the following formula for estimating a single population proportion was used:

$$n = \frac{(z_{1-\alpha/2})^2 \, p(1-p)}{d^2}$$

Where:

n = required sample size

Z = standard normal deviate corresponding to the desired confidence level (e.g., 1.96 for 95%)

p = expected prevalence (0.3)

d = margin of error (commonly set at 0.05)

$$n = \frac{(1.96)^2 \cdot 0.3(1-0.3)}{(0.05)^2}, n = 322.68$$

n≈ 323 participants

Therefore, the minimum required sample size is approximately 323 participants. To account for potential data loss or incomplete samples (e.g., due to PCR failure or poor DNA quality), an additional 10% were added: ^{n}adjusted =323+(323×0.10)=355.3≈355. Thus, a total of 355 malaria-positive samples were targeted for inclusion in the study.

3.3.2. Participant Recruitment and Sample Processing

In this study, malaria case detection will begin at the community level where Community Health Workers (CHWs) will perform rapid diagnostic tests (RDTs) to screen individuals for malaria infection. Those who test positive and provide informed consent will be referred to the nearest health center for further clinical assessment and biological sample collection. At the health center, and prior to the administration of any malaria treatment, venous blood samples will be collected into EDTA tubes by trained laboratory technicians. In addition, thick and thin blood smears will be prepared and stained using Giemsa to allow for microscopic identification of Plasmodium species. For cases confirmed as *Plasmodium falciparum*, the EDTA blood samples will be forwarded to the National Reference Laboratory (NRL) for advanced molecular analysis. All collected samples and slides will be packaged according to strict biosecurity standards, including the use of triple packaging and maintenance of the cold chain. The samples will be initially transported to the District Hospital and then forwarded within 24 hours to the NRL to preserve sample integrity. At the NRL, DNA was extracted from the EDTA blood samples. Molecular testing involved polymerase chain reaction (PCR) and next generation sequencing(Nanopore). The primary focus of sequencing will be the *Pfkelch13* (PfK13) gene, known for its association with artemisinin resistance. Additional genetic markers may also be sequenced to investigate parasite genetic relatedness and clustering patterns.

3.3.3. Training of Data Collectors

A structured and comprehensive three-day training program was implemented for Community Health Workers (CHWs), nurses, and laboratory personnel directly involved in participant recruitment, clinical assessment, biological sample collection, and data documentation. This preparatory training, conducted prior to the commencement of data collection, was designed to ensure rigorous adherence to the study protocol and promote procedural uniformity across all study sites. Core components of the training included standardized methodologies for malaria case identification, informed consent acquisition, and the appropriate clinical referral of confirmed cases. Laboratory personnel received specialized instruction in venous blood collection, preparation of both thick and thin blood smears, and the secure transport of specimens using triple-layer packaging in accordance with biosecurity regulations. Additionally, data collectors were thoroughly oriented on the study protocol, ethical research practices, and strategies to safeguard participant confidentiality. All personnel were also trained in the accurate and consistent use of study tools, including data collection forms and participant questionnaires, to ensure high-quality data capture and harmonized implementation throughout the study.

3.3.4. Study Variables

This study examines the molecular and epidemiological factors associated with genetic polymorphisms in Plasmodium falciparum, with a focus on variations in the AMA1 and PfK13 genes. The dependent variables include the presence, absence, and type of genetic mutations, particularly SNPs and amino acid substitutions due to their relevance to vaccine-targeted antigenic variation (AMA1) and artemisinin resistance (PfK13).

The independent variables fall into four categories. Socio-demographic factors, such as age, sex, education, and occupation may influence exposure and health-seeking behaviour. Environmental and household factors include household size, housing conditions, malaria prevention practices (e.g., net use, indoor residual spraying), and access to health information. Behavioral and healthcare access variables encompass travel history, contact with community health workers, proximity to health facilities, and treatment behaviors factors relevant to transmission and potential resistance spread. Clinical and parasitological variables, such as fever history, parasitemia level, hemoglobin concentration, and recent treatment, provide insights into disease severity and host-pathogen interactions.

Together, these variables provide a multidimensional framework for evaluating the genetic diversity, molecular clustering, and geographic distribution of *P. falciparum* infections in Rwanda, supporting the study's broader goal of informing vaccine development and artemisinin resistance surveillance.

3.4. Sample Collection and Processing

A total of 200 µL of venous blood was collected from each participant into EDTA anticoagulant tubes. For initial microscopic confirmation, both thick and thin blood smears were prepared on-site and stained using Giemsa stain. After smear preparation, all blood samples were stored at 4°C and subsequently transported to the National Reference Laboratory (NRL) under strict cold chain conditions to preserve sample integrity for downstream molecular analysis.

3.5. Genomic DNA Extraction

Genomic DNA was extracted from whole blood using the Qiagen QIAamp DNA Mini Kit according to the manufacturer's instructions. In brief, 200 µL of each blood sample was mixed with 20 µL of Proteinase K and 200 µL of AL lysis buffer, followed by incubation at 56°C for 10 minutes to promote complete cell lysis. Subsequently, 200 µL of ethanol was added to the lysate, and the mixture was transferred to a silica membrane spin column to facilitate DNA binding. The columns were washed sequentially with AW1 and AW2 wash buffers to eliminate impurities. Finally, purified DNA was eluted using 200 µL of AE buffer and stored at −20°C for subsequent molecular and bioinformatics analyses.

3.6. Whole Genome Amplification (WGA)

To ensure sufficient DNA input for downstream analyses, Selective Whole Genome Amplification (sWGA) was performed on samples exhibiting low parasitemia levels (<1,000 parasites/µL). This was achieved using the EquiPhi29 DNA Polymerase protocol as adapted from the NOMADS MVP study. The amplification process began with DNA denaturation followed by primer annealing using a pool of ten degenerate primers (Pf1–Pf10) specifically designed to selectively target *Plasmodium falciparum* genomic regions. The reaction was incubated at 45°C for 60 minutes to allow for optimal primer binding and strand displacement activity, followed by an extension phase at 65°C for 10 minutes. Amplification products were subsequently diluted at a ratio of 1:20 and quantified using the Qubit fluorometric system. Only successfully amplified DNA products meeting the required concentration thresholds were retained for downstream gene-targeted PCR and sequencing workflows.

3.7. Molecular Amplification and Sequencing

3.7.1. Target Gene Selection

Two *P. falciparum* genes were selected for molecular characterization in this study: PfK13 and PfAMA1. The PfK13 gene was chosen for its well-established role in mediating artemisinin resistance, with particular attention paid to non-synonymous mutations within its propeller domain. In contrast, PfAMA1 served as a molecular marker for assessing antigenic diversity and population structure, with amplification focused on Domain I and II regions known for their immunogenic relevance and high polymorphism.

3.7.2. Polymerase Chain Reaction (PCR) Amplification

We performed Polymerase Chain Reaction (PCR) to amplify two *Plasmodium falciparum* genes: the kelch 13 (*PfK13*) propeller domain, associated with artemisinin resistance, and the apical membrane antigen 1 (*PfAMA1*) gene, a highly polymorphic blood-stage vaccine candidate. *PfK13* amplification employed the forward primer 5′-CGGAGTGACCAAATCTGGGA-3′ and reverse primer 5′-GCCTTGTTGAAAGAAGCAGA-3′, adapted from Ariey et al. (2014), using Solis BioDyne reagents. *PfAMA1* amplification targeted immunogenic domains I and II with validated Malaria Vaccine Program (MVP) primer sets described by Hodder et al. (2001). Each 25 µL reaction contained 2× KAPA HiFi ReadyMix, gene-specific primers, and 20–50 ng of template DNA. Thermal cycling comprised initial denaturation at 95 °C for 3 minutes, 35 cycles of 98 °C for 20 seconds and 60 °C for 3 minutes, and a final extension at 60 °C for 10 minutes.

3.7.3. Amplicon Verification and Purification

The success of PCR amplification was verified through 1.5% agarose gel electrophoresis, enabling visual confirmation of amplicon integrity and expected fragment sizes. PCR products were subsequently purified using AMPure XP magnetic beads according to the manufacturer's guidelines. This step was critical for removing primer dimers, residual primers, and non-specific amplification products, thereby ensuring clean templates for high-quality sequencing.

3.7.4. DNA Sequencing

Purified PCR amplicons were subjected to long-read sequencing using the **Oxford Nanopore Technologies (ONT)** MinION platform. ONT sequencing was selected for its ability to generate long

contiguous reads, which is particularly advantageous for accurate haplotype reconstruction, detection of structural variants, and resolving complex genetic regions within both **PfK13** and **PfAMA1** loci.

3.8. Bioinformatics and Genetic Data Analysis

3.8.1. Raw Data Preprocessing

The raw sequencing reads, initially generated in FASTQ format, underwent a systematic quality control and preprocessing workflow to ensure the integrity and reliability of downstream analyses. Quality assessment was first performed using FastQC, which provided detailed visual reports on base quality scores, GC content distribution, and sequence duplication levels, thereby identifying potential biases or technical artefacts. Following quality assessment, adapter sequences and low-quality bases were removed using Trimmomatic, employing a sliding window approach and base quality thresholds to retain high-confidence reads.

Subsequently, the cleaned reads were processed through alignment and multiple sequence alignment (MSA) strategies tailored to the target gene and sequencing platform. For PfK13, a relatively conserved gene with shorter amplicon lengths, the high-quality reads were mapped to the Plasmodium falciparum reference genome using the Burrows-Wheeler Aligner (BWA), a robust tool for efficient short-read alignment. In contrast, due to the high sequence variability and antigenic diversity characteristic of PfAMA1, multiple sequence alignments were constructed using MAFFT, which is particularly effective for generating accurate alignments across highly polymorphic regions. This dual-alignment strategy ensured optimal handling of both conserved and variable loci, forming a robust foundation for subsequent variant calling, haplotype reconstruction, and evolutionary analysis.

3.8.2. SNP Calling and Variant Annotation

After completing sequence alignment, the analysis proceeded with the identification of single nucleotide polymorphisms (SNPs) using the SAMtools and BCFtools pipeline. This integrated approach enabled accurate variant detection across aligned reads by generating binary variant call format (BCF) files from alignment data and applying probabilistic models to infer base-level variants.

To improve variant reliability, all identified SNPs were manually curated and visually inspected using Geneious. This step ensured the validation of ambiguous base calls and the exclusion of sequencing artefacts, misalignments, or low-confidence positions. In parallel, all multiple sequence alignments

particularly those generated for the highly polymorphic PfAMA1 gene were manually reviewed and corrected using AliView. This review process allowed the removal of spurious insertions and enforced codon-aware alignment, preserving reading frames and preventing translation disruptions.

The workflow then applied variant annotation using BioPython, which enabled the functional interpretation of mutations based on curated gene models and reference protein databases. Annotation focused on identifying the mutation type (e.g., synonymous or non-synonymous), amino acid changes, and positions within known functional or antigenic domains. This annotation step provided critical insights into the potential biological significance of observed variants, facilitating downstream analyses related to haplotype clustering, evolutionary selection, and vaccine or drug resistance implications.

3.8.3. Amino Acid Translation and Domain Mapping

Identified SNPs were translated in-frame to determine their corresponding amino acid substitutions. These amino acid changes were then mapped to annotated functional domains within *PfK13* (e.g., propeller domain) and *PfAMA1* (e.g., Domain I and II). This domain-level annotation allowed for targeted exploration of the functional relevance of observed polymorphisms in the context of immune evasion or drug resistance mechanisms.

3.8.4. Haplotype Construction and Visualization

Haplotypes were reconstructed from nonsynonymous SNPs across the *PfK13* and *PfAMA1* gene regions. This process focused specifically on amino acid-changing variants associated with either artemisinin resistance (for *PfK13*) or antigenic variation (for *PfAMA1*). Haplotypes were subsequently visualized using network-based methods and sequence clustering tools, enabling molecular characterization of circulating *P. falciparum* lineages.

3.8.5. Genetic Diversity and Neutrality Metrics

To evaluate genetic diversity and infer selective pressures acting on target genes, neutrality statistics including **Tajima's D** and **Fu and Li's D** were calculated. These metrics were computed using custom **Python scripts**, providing insights into demographic history and the potential presence of balancing or directional selection. Diversity metrics were assessed across different regions of each gene to identify hotspots of evolutionary constraint or adaptation.

3.8.6. Linkage Disequilibrium (LD) and Principal Component Analysis (PCA)

Pairwise linkage disequilibrium (LD) analysis was conducted using the scikit-allel Python package, with r^2 values calculated to assess non-random associations between SNP loci within each gene. To explore genetic structure and potential population stratification, Principal Component Analysis (PCA) was performed on the resulting biallelic SNP matrix using scikit-learn. PCA plots facilitated the visualization of genetic clustering patterns among isolates from Huye and Kirehe, potentially reflecting transmission dynamics or region-specific selection.

3.8.7. Phylogenetic and Network Analysis

Phylogenetic relationships among parasite isolates were inferred by constructing Neighbour-Joining (NJ) trees using MEGA software based on aligned *PfK13* and *PfAMA1* sequences. These trees provided insight into evolutionary relatedness and lineage divergence, particularly in the context of known resistance-associated haplotypes. Complementary haplotype network analyses were also performed to visualize mutational connections between closely related isolates.

3.8.8. Co-occurrence Analysis of PfK13 and PfAMA1 Variants

To examine the potential interaction between antigenic variation and drug resistance, co-occurrence analysis was conducted to evaluate associations between *PfK13* mutations and *PfAMA1* haplotype clusters. This integrated analysis aimed to determine whether specific resistant genotypes are enriched within distinct antigenic backgrounds, which may have implications for both transmission dynamics and vaccine design.

3.9. Data Analysis

To investigate molecular clustering and transmission patterns of Plasmodium falciparum in Rwanda, genetic data were analyzed through a combination of phylogenetic, haplotype, and population genetics approaches. Clustering analysis was performed using sequence similarity-based methods, including phylogenetic tree reconstruction and haplotype network generation, to infer evolutionary relationships among AMA1 variants and to identify genetically distinct parasite lineages.

These analyses were used to characterize haplotype diversity and structure within and between sampling sites, Phylogenetic trees were constructed using the Neighbor-Joining method, while haplotype networks

were generated using frequency-based algorithms to visualize connections among allelic variants and infer transmission dynamics.

In parallel, the analysis assessed the prevalence and geographic distribution of PfK13 mutations, with a specific focus on known artemisinin resistance markers such as R561H. Statistical comparisons were conducted to evaluate associations between PfK13 variants, AMA1-defined clusters, and regional origin (Huye vs Kirehe).

3.10. Ethical Considerations

This study received ethical approval under the framework of the ARMEA Project, following all applicable institutional and national research ethics guidelines. Participation was entirely voluntary, and written informed consent was obtained from all participants or their legal guardians prior to sample collection. To ensure participant confidentiality, all personal identifiers were removed, and unique coded identifiers were assigned to biological samples and associated data. All data handling, storage, and analysis procedures strictly adhered to protocols designed to protect participant privacy, data security, and ethical research conduct.

CHAPTER 4: RESULTS AND DISCUSSION

This study analyzed genetic variation in two *Plasmodium falciparum* genes, *PfAMA1* and *PfK13*, from clinical isolates collected in Huye and Kirehe districts, Rwanda. The *PfAMA1* gene, which encodes a key blood-stage antigen and leading malaria vaccine candidate, showed substantial haplotype diversity, particularly within immune-exposed regions. Similar high levels of polymorphism have been reported in African and Asian parasite populations (Barry et al., 2009; Hodder et al., 2001; Kusi et al., 2017; Tetteh et al., 2008), and are thought to result from balancing selection driven by host immune pressure (Dutta et al., 2005; World Health Organization, 2023b). This genetic variability poses challenges for single-allele vaccine designs, supporting the need for multi-allelic or conserved-region approaches to maintain vaccine efficacy (Dutta et al., 2009; Ouattara et al., 2020).

This chapter presents the findings on the genetic diversity and evolutionary dynamics of the *Plasmodium falciparum* PfAMA1 and PfK13 genes, based on clinical isolates collected from Huye and Kirehe districts in Rwanda. The analysis focuses on characterizing polymorphisms, haplotype structures, and resistance-associated mutations, with the aim of elucidating regional patterns of parasite population structure, artemisinin resistance, and implications for targeted vaccine development.

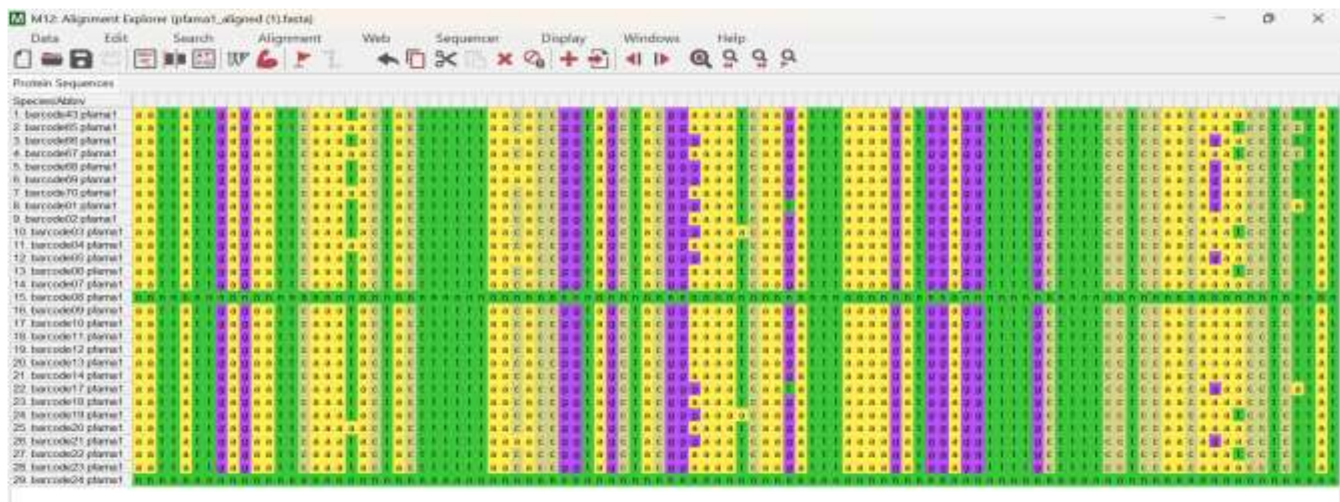


Figure 1: *shows the multiple sequence alignment of the Plasmodium falciparum AMA1 gene across 29 barcoded clinical isolates from Rwanda. Conserved nucleotide sites are represented in green, while polymorphic loci appear in variant colors (e.g., purple, yellow), indicating nucleotide substitutions at specific genomic positions.*

The visualized alignment (Figure 1) clearly illustrates extensive polymorphism across the AMA1 coding region. Numerous colored columns reflect point mutations among isolates, with hotspots of variation concentrated within regions known to encode antigenic epitopes. These substitutions likely reflect immune-driven selection, as AMA1 is a surface-expressed protein targeted by host antibodies. The observed diversity supports the presence of multiple haplotypes and provides molecular evidence for localized evolutionary pressures in high-transmission areas. This aligns with **Objective 1**, confirming genetic polymorphism and informing later clustering analysis.

Figure 2: *presents the multiple sequence alignment of the PfK13 gene across the same 29 P. falciparum isolates. Highly conserved regions are displayed in green, while variable loci, especially within the propeller domain, are highlighted in other colors.*

The PfK13 alignment (Figure2) reveals fewer polymorphic sites compared to AMA1, emphasizing its conserved functional role. However, variation is notably concentrated in the C-terminal propeller region, which has been implicated in artemisinin resistance. The visual distribution of substitutions suggests that while the gene is under functional constraint, specific mutations likely adaptive emerge under drug pressure. This is central to **Objective 3**, where mutations within PfK13 are linked to resistance phenotypes and spatial patterns.

Mutation Type	Amino Acid Change	gt	Ref Allele	Altenative Alle	Origin(District)	Mutation Count	Frequency %
missense	G449A	1/1	C	G	HUYE	2	6.90%

missense	N490T	0/1	T	G	KIREHE	1	3.40%
missense	R561H	0/1	C	T	KIREHE	3	10.3%

Table 1**: (PfK13):** *Mutations in the PfK13 gene from 29 P. falciparum samples in Huye and Kirehe. R561H, a proven resistance marker, appeared only in Kirehe (10.3%, mixed form). G449A (6.9%, fixed form) was found only in Huye, and N490T (3.4%, mixed form) only in Kirehe. These patterns show possible local spread of resistance in Kirehe and unique low-frequency changes in each district.*

Mutation Type	Amino Acid Change	Genotype	Allele	Allele	Origin (District)	Mutation Count	Frequency (%)
missense	D196N	0/1	G	A	Huye	1	3.4
missense	D204N	0/1	G	A	Huye	1	3.4
missense	D244N	0/1	G	A	Kirehe	1	3.4
missense	E187N	0/1	G	A	Huye	3	10.3
missense	E197G	0/1	A	G	Huye	1	3.4
missense	E197Q	0/1	G	C	Huye	1	3.4
missense	E267Q	0/1	G	C	Huye	1	3.4
missense	E267Q	0/1	G	C	Kirehe	1	3.4
missense	F201L	0/1	T	A	Huye	1	3.4
missense	G172E	0/1	G	A	Huye	1	3.4
missense	G172E	0/1	G	A	Kirehe	1	3.4
missense	H200D	0/1	C	G	Huye	1	3.4
missense	I225N	0/1	T	A	Kirehe	1	3.4
missense	I282K	0/1	T	A	Kirehe	1	3.4
missense	K206E	0/1	A	G	Huye	2	6.9
missense	K269I	0/1	A	T	Huye	1	3.4
missense	L189H	0/1	T	A	Huye	1	3.4
missense	M190I	0/1	G	A	Huye	1	3.4
missense	N162K	0/1	T	A	Kirehe	1	3.4
missense	N173K	0/1	T	A	Huye	1	3.4
missense	Y175D	0/1	T	G	Huye	2	6.9

missense	Y175D	0/1	T	G	Kirehe	1	3.4
missense	Y207D	0/1	T	G	Huye	1	3.4

Table 2: (PfAMA1): Mutations in the PfAMA1 gene from Huye and Kirehe. A total of 24 amino acid changes were found, most rare (3.4%), with some recurring—E187N (10.3%), K206E (6.9%), and Y175D (6.9%). Most were in mixed form and showed a mix of shared and district-specific patterns, indicating immune-driven diversity and regional parasite differences.

We analysed genetic variation in two key *Plasmodium falciparum* genes: PfK13, which is used to detect signs of artemisinin resistance, and *PfAMA1*, an important blood-stage vaccine candidate. These analyses were based on 29 clinical isolates collected from malaria patients in Huye and Kirehe districts, the two main study sites in this research.

For *PfK13,* three amino acid changes were identified: R561H, G449A, and N490T. The R561H mutation was found in three isolates (10.3%), all from Kirehe, and appeared in a mixed form (heterozygous), meaning the infections contained both mutant and wild-type parasites. This mutation is a World Health Organization–confirmed marker of artemisinin resistance and has been associated with slower parasite clearance. Its presence in only one district suggests that resistant parasites may have recently emerged there or been introduced from elsewhere, and are now beginning to spread. The mixed form indicates that resistance may be circulating within parasite populations without fully replacing sensitive strains, a situation that can allow resistant types to persist and expand over time. G449A was observed in two isolates (6.9%), only in Huye, and was present in a fixed form (homozygous), meaning all parasites in those infections carried the change—possibly reflecting stability or local adaptation in that setting. N490T appeared in one isolate (3.4%), only in Kirehe, and was mixed. Although G449A and N490T are not recognised markers of resistance, their occurrence in specific districts points to differences in the parasite populations between regions, which could become more important if drug pressure continues to act on these variants.

For PfAMA1, the level of variation was much higher, with 24 different amino acid changes identified. All were missense mutations, meaning they altered the protein sequence. Most were rare, appearing in just one sample (3.4%), but several were more frequent, such as E187N (10.3%), K206E (6.9%), and Y175D (6.9%). In most cases, these changes appeared in a mixed form, suggesting that patients were often infected with more than one parasite type at the same time. Some mutations were shared between

the two districts, such as E267Q and G172E, while others were restricted to one location, such as D244N in Kirehe and N173K in Huye. This mix of common and unique mutations, combined with the large number of rare variants, shows that the parasite maintains a wide variety of PfAMA1 forms within and between districts. Such variety is important for the parasite because it helps avoid detection by the human immune system, making it harder for naturally acquired immunity—or vaccines— to provide lasting protection. These findings suggest that a PfAMA1-based vaccine designed for Rwanda would need to account for multiple variants to achieve broader effectiveness.

Taken together, these results show two very different patterns: PfK13 had fewer changes, but one of them (R561H) has major significance for drug resistance and was concentrated in one district. PfAMA1 displayed broad diversity across both districts, reflecting the parasite's long-term strategy to survive immune pressure. These findings directly support the study objectives by showing that K13 variation can be used to detect possible resistance hotspots, while AMA1 diversity needs to be mapped and monitored to guide vaccine development and predict how well a vaccine might work in different regions of Rwanda.

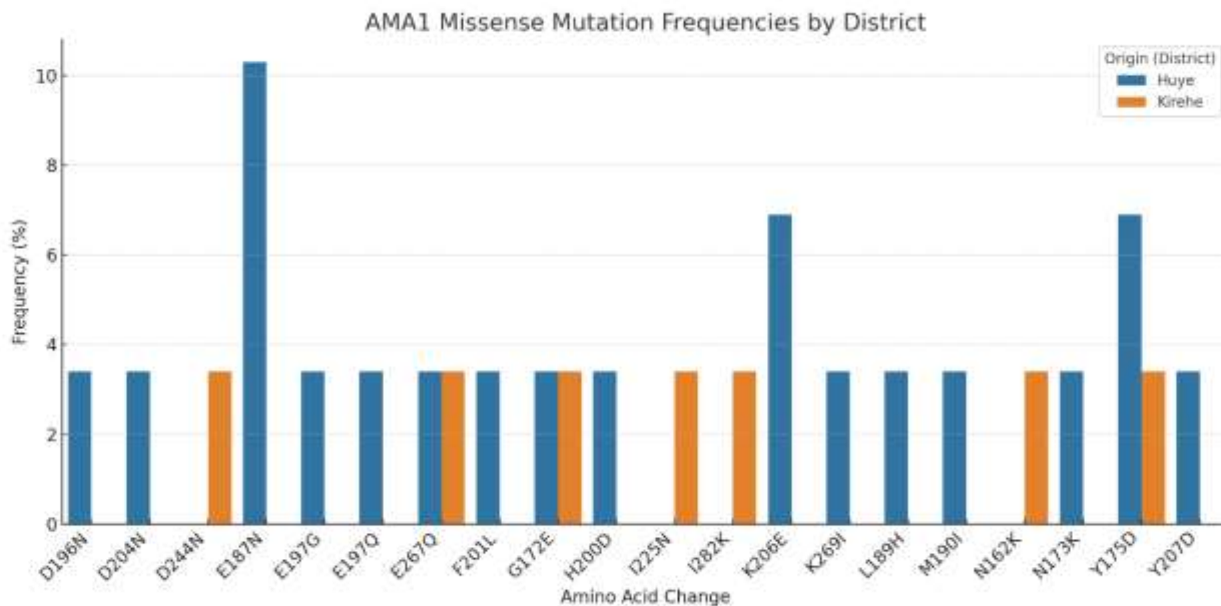


Figure 3: This bar chart presents the frequency distribution of missense mutations in the PfAMA1 gene across 29 Plasmodium falciparum isolates. Each bar represents the percentage of isolates with a specific

amino acid substitution, separated by district of origin Huye (blue) and Kirehe (orange). Frequencies are calculated as a percentage of the total sample size (n=29).

The bar chart highlights the distribution and frequency of non-synonymous mutations in the AMA1 gene across the two study sites. The mutations E187N, K206E, and Y175D appear most frequently, with E187N reaching 10.3%, and K206E and Y175D each at 6.9%. These mutations occur exclusively or predominantly in Huye, suggesting a localized expansion of particular haplotypes likely shaped by immune-driven selection pressure.

Mutations such as E267Q, G172E, and Y175D appear in both districts, indicating potential gene flow or the circulation of shared parasite lineages across regions. The presence of low-frequency mutations (3.4%) across multiple positions also underscores the high allelic diversity of the AMA1 gene in this population.

Overall, Huye exhibits a higher mutational burden compared to Kirehe, supporting the idea that Huye may be a hotspot of antigenic variation. This aligns with findings from the alignment tables and heatmap, and reinforces the need to consider region-specific AMA1 haplotypes in vaccine development. These results directly support Objective 1 of the study, which focuses on characterizing the genetic polymorphism and haplotype diversity of the AMA1 gene in Rwanda.

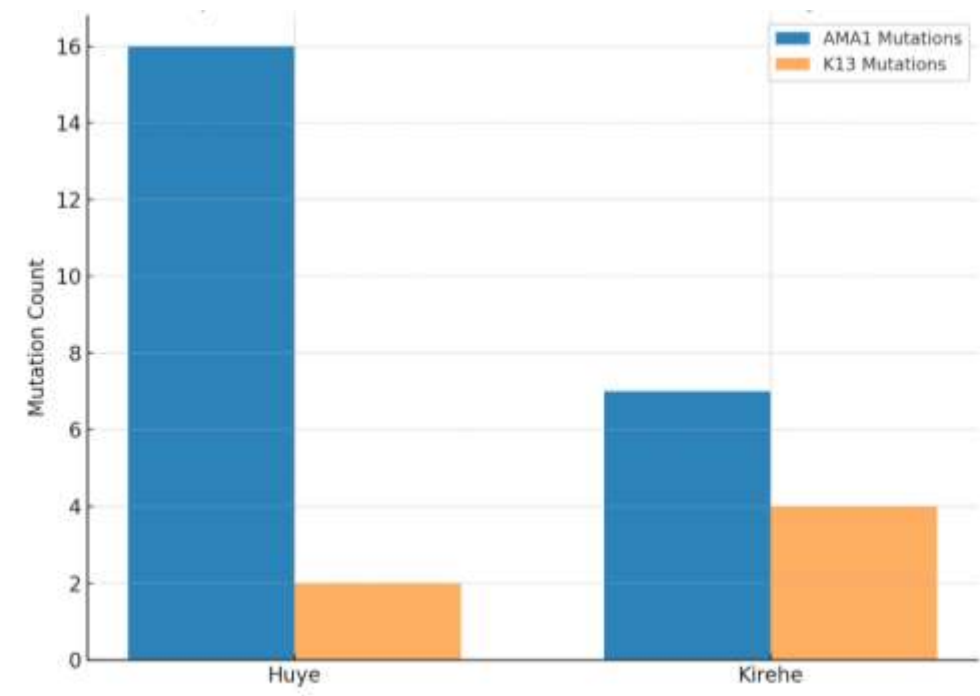


Figure 4: This grouped bar chart compares the total number of missense mutations identified in the PfAMA1 and PfK13 genes across clinical isolates from Huye and Kirehe districts. The height of each bar represents the count of unFigure 4ique mutations detected in each gene per district, based on 29 samples.

When we compared the number of AMA1 and K13 gene mutations found in the Huye and Kirehe districts, we saw clear differences in both the total number and the way these mutations were spread out geographically. AMA1 mutations were much more common overall, with Huye showing 16 unique mutations and Kirehe showing 7. This supports earlier results suggesting that the AMA1 gene is highly variable, especially in Huye. Since AMA1 plays a key role in helping malaria parasites enter red blood cells, this high level of variation might help the parasite avoid being attacked by the immune system in that region. It also suggests that parasites in Huye might be evolving quickly in response to host immunity.

On the other hand, K13 mutations were less frequent but more focused in Kirehe. This includes the R561H mutation, which is already known to be linked to resistance to artemisinin the main drug used to treat malaria. R561H was found only in Kirehe, along with another potential resistance mutation (N490T). Huye had only two K13 mutations (G449A), and none were confirmed resistance markers. This difference suggests that Kirehe might be a hotspot for the development or spread of drug-resistant parasites, possibly due to how antimalarial drugs are used or other local factors.

These findings strongly support two key goals of the study. First, they confirm that AMA1 shows a lot of genetic variation (Objective 1), especially in Huye. Second, they show that resistance-related K13 mutations are not spread evenly but are mostly found in Kirehe (Objective 3). This highlights the importance of doing region-specific monitoring and using targeted approaches when designing vaccines or tracking drug resistance. In short, not all areas face the same malaria challenges, and strategies should reflect local parasite genetics.

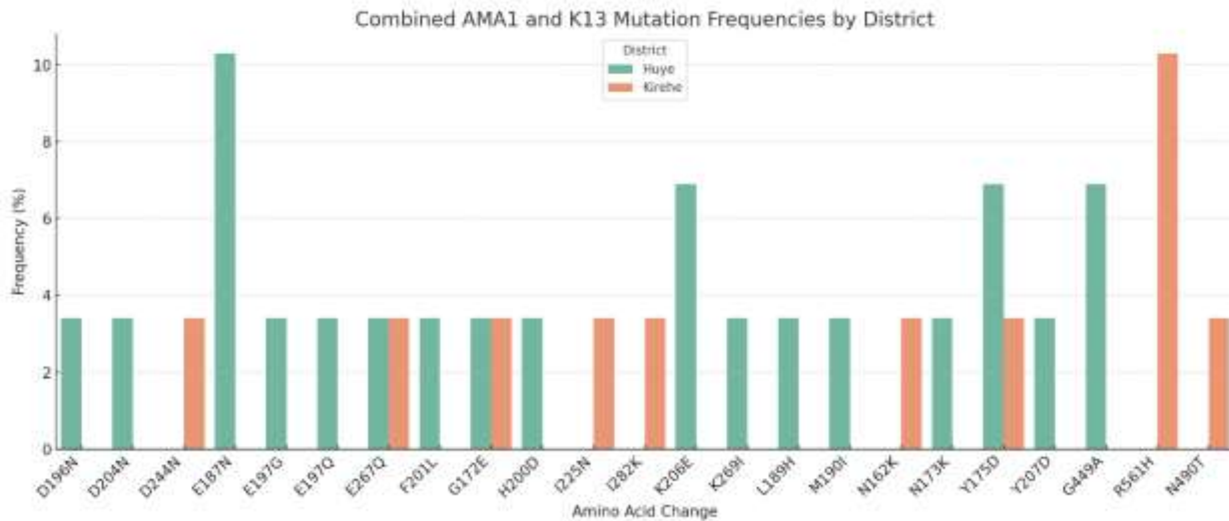


Figure 5: Distribution of AMA1 and K13 Missense Mutations in P. falciparum Clinical Isolates from Huye and Kirehe Districts, Rwanda

This figure shows how malaria parasites are changing (mutating) in two districts of Rwanda Huye and Kirehe based on the presence of specific mutations in two important genes: AMA1 and K13. The chart compares how often each mutation occurred in samples from each district, helping us understand how these changes may affect both malaria treatment and vaccine development.

Most of the mutations were found in the AMA1 gene, which helps the parasite invade red blood cells and is often targeted by vaccines. These AMA1 mutations appeared mostly in samples from Huye. One mutation, E187N, occurred most frequently (over 10%) and was found only in Huye. Other mutations such as Y175D and K206E were also mainly seen in Huye. This suggests that in Huye, the parasite may be evolving under immune pressure trying to escape the immune system's response. This is an important observation when considering which vaccine strategies might work best in that region.

In contrast, the K13 gene mutations were found only in samples from Kirehe. The most important among them, R561H, is officially recognized by the World Health Organization as a marker of artemisinin resistance, which is the main drug used to treat malaria. Other K13 mutations like G449A and N490T were also present in Kirehe and may represent early signs of drug resistance. These findings suggest that Kirehe may be a hotspot for drug-resistant malaria, which could lead to treatment failure if not carefully monitored.

In summary, the chart shows that Huye has more AMA1 mutations, which points to immune system pressure, while Kirehe has more K13 mutations, which may reflect drug pressure from artemisinin use. These differences mean malaria control strategies should be adjusted to each district's situation, improving vaccines for Huye and focusing on drug resistance surveillance in Kirehe.

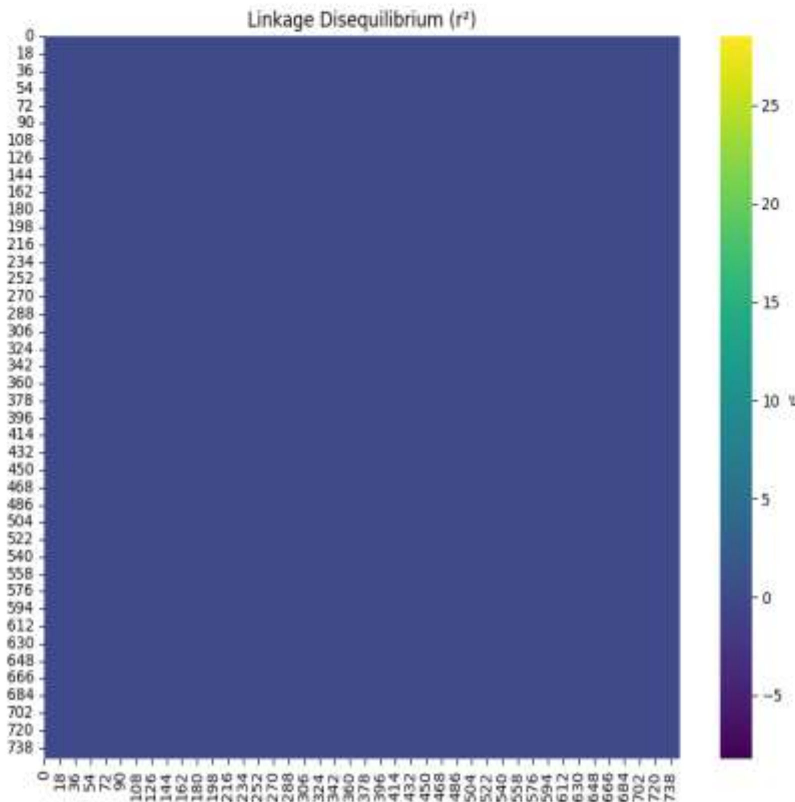


Figure 6: Linkage Disequilibrium Pattern Across the AMA1 Gene (r^2 Heatmap).

This figure displays a heatmap of pairwise linkage disequilibrium (LD), measured by $\mathbf{r^2}$, across SNPs in the AMA1 gene of *Plasmodium falciparum*. Each axis (X and Y) represents specific nucleotide positions along the gene, while the color gradient indicates the strength of association between SNPs at those positions. In this color scale, dark blue represents low LD (r^2 near 0), meaning weak or no correlation between variants, while yellow would represent high LD (r^2 close to 1), indicating strong correlation and co-inheritance.

The heatmap is predominantly dark blue, showing that most SNP pairs have very low LD, suggesting that these mutations are not inherited together frequently. This pattern reveals the absence of strong LD blocks, a typical sign of extensive recombination or high genetic diversity, both of which break down linkage over time. Furthermore, the lack of yellow or red zones means that there are no dominant haplotypes or strongly linked mutation pairs within the AMA1 gene in this population.

Scientifically, this pattern supports the idea that the AMA1 gene is under strong balancing selection likely due to host immune pressure. In response, the parasite accumulates diverse, unlinked mutations to evade immune detection. This polymorphic nature of AMA1 has direct implications for vaccine design, as it implies that multiple independent haplotypes co-exist, making it harder for a single antigen to provide broad protection. The observed LD profile also indicates rapid parasite evolution, aligning with findings from other regions where AMA1 diversity undermines vaccine efficacy.

In conclusion, this LD heatmap reflects a genetically diverse, highly recombining AMA1 locus, which not only complicates vaccine formulation but also emphasizes the adaptive capabilities of *P. falciparum* in the face of host immune responses.

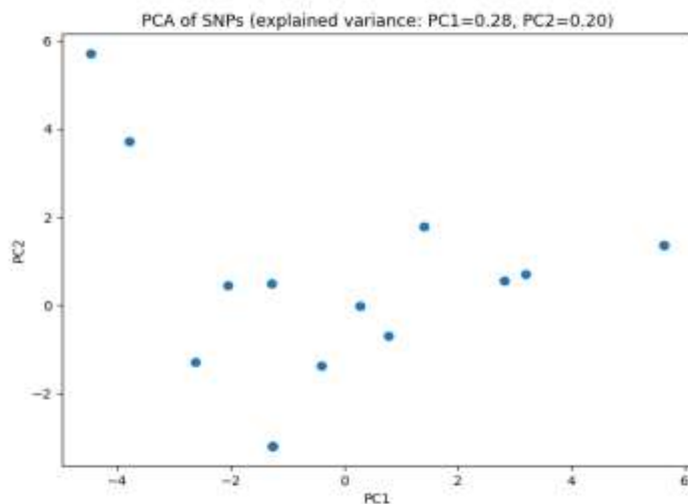


Figure 7: Linkage Disequilibrium Pattern Across the AMA1 Gene (r^2 Heatmap).

This scatter plot shows the results of Principal Component Analysis (PCA) conducted on single nucleotide polymorphisms (SNPs) from the AMA1 and K13 genes of Plasmodium falciparum. Each dot represents a single parasite sample. The position of each dot is determined by its genetic variation,

summarized into two main axes: PC1 (x-axis) and PC2 (y-axis). These two components together explain 48% of the total genetic differences observed between samples (PC1 = 28%, PC2 = 20%).

The spread of the dots across the plot shows that the malaria parasites from Rwanda are genetically diverse meaning there are many differences in their DNA. This diversity is expected in genes like AMA1, which are under pressure from the human immune system and from drug treatments. The fact that the dots don't cluster tightly or show clear groups means that parasite populations from different districts (like Huye and Kirehe) are not genetically separated. This suggests that there is ongoing mixing of parasites between districts, likely due to human movement or shared transmission zones.

However, some samples appear far away from the others these are called outliers. These outliers could be carrying rare or unusual genetic mutations, such as R561H in the K13 gene, which is known to be associated with artemisinin resistance. The presence of such outliers highlights the need to monitor these unusual strains closely, as they may spread or evolve differently.

This PCA confirms that clinical malaria infections in Rwanda are caused by a genetically diverse population of parasites. The lack of strong geographical structure means that interventions (like vaccines or drugs) must account for shared risk across regions. Meanwhile, the presence of distinct outlier genotypes suggests early signs of local adaptation or emerging resistance, which supports the study's aim to identify molecular clusters and evolutionary dynamics.

Gene	Number of Sequences (m)	Number of Segregating Sites (S)	Segregating Site Proportion (ps)	Theta (θ)	Pi (π)	Tajima's D
AMA1	29	2494	0.607996	0.154818	0.200371	1.158407
PfK13	29	2626	0.801098	0.203989	0.200304	-0.07112

Table 3: Summary of Genetic Diversity and Neutrality Tests for AMA1 and PfK13 Genes

S: Number of sites where the DNA varies across samples.

ps: Proportion of the gene that has variation.

θ (Theta): Expected variation based on mutation rate.

π (Pi): Observed genetic variation.

Tajima's D: Test for selection. Positive means balancing selection (multiple variants kept), negative means recent expansion or purifying selection

This summary shows key genetic diversity results for the AMA1 and PfK13 genes based on 29 samples of *Plasmodium falciparum.* It includes important population genetics values such as the number and proportion of variable sites (segregating sites), nucleotide diversity (π), Watterson's theta (θ), and Tajima's D. These help us understand how much genetic variation exists and what evolutionary pressures might be influencing these genes.

Both genes have a large number of variable positions 2,494 in AMA1 and 2,626 in PfK13 which confirms they are highly polymorphic. However, the proportion of variation is higher in PfK13 (80%) than in AMA1 (61%), suggesting that PfK13 has more changes across its length. This could be due to stronger or more recent selection pressures, such as drug use.

Nucleotide diversity (π), which looks at the average differences between gene sequences, is almost the same for both genes (~0.200), showing similar overall variation at the base-pair level.

Theoretical diversity (θ) is lower in AMA1 (0.15) than in PfK13 (0.20), which means that the observed variation in AMA1 is greater than what is expected under neutral evolution. This is supported by the positive Tajima's D value (1.15) for AMA1, which indicates **balancing selection** a process that keeps different variants in the population. This is typical for genes like AMA1 that are involved in immune evasion and vaccine response.

PfK13, on the other hand, has a Tajima's D close to zero (-0.07), which suggests that it is evolving mostly neutrally or under weak purifying selection. Even though it has many mutations, these changes might be recent and haven't yet become common, like the known resistance mutation R561H.

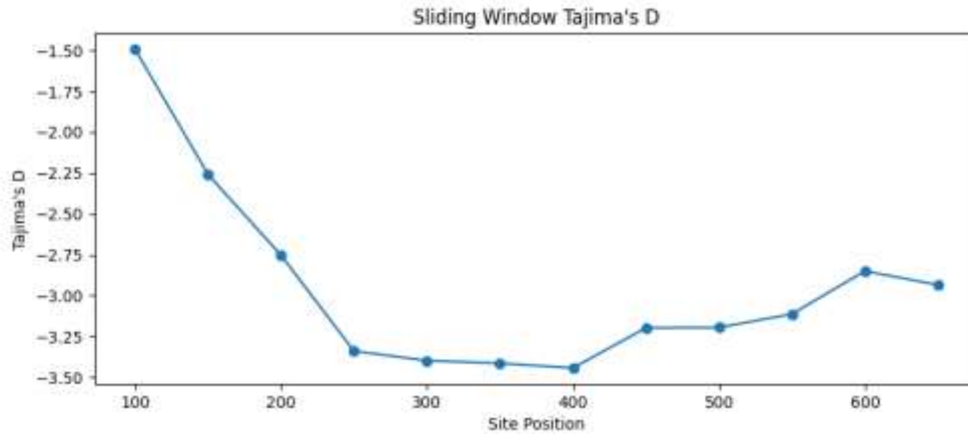


Figure 8: Sliding Window Analysis of Tajima's D across the *AMA1* Gene in *Plasmodium falciparum.*

This figure shows how genetic variation changes across the *AMA1* gene by plotting Tajima's D values using a sliding window approach. Each point represents a genomic window, showing whether mutations in that region are evolving neutrally or under selection.

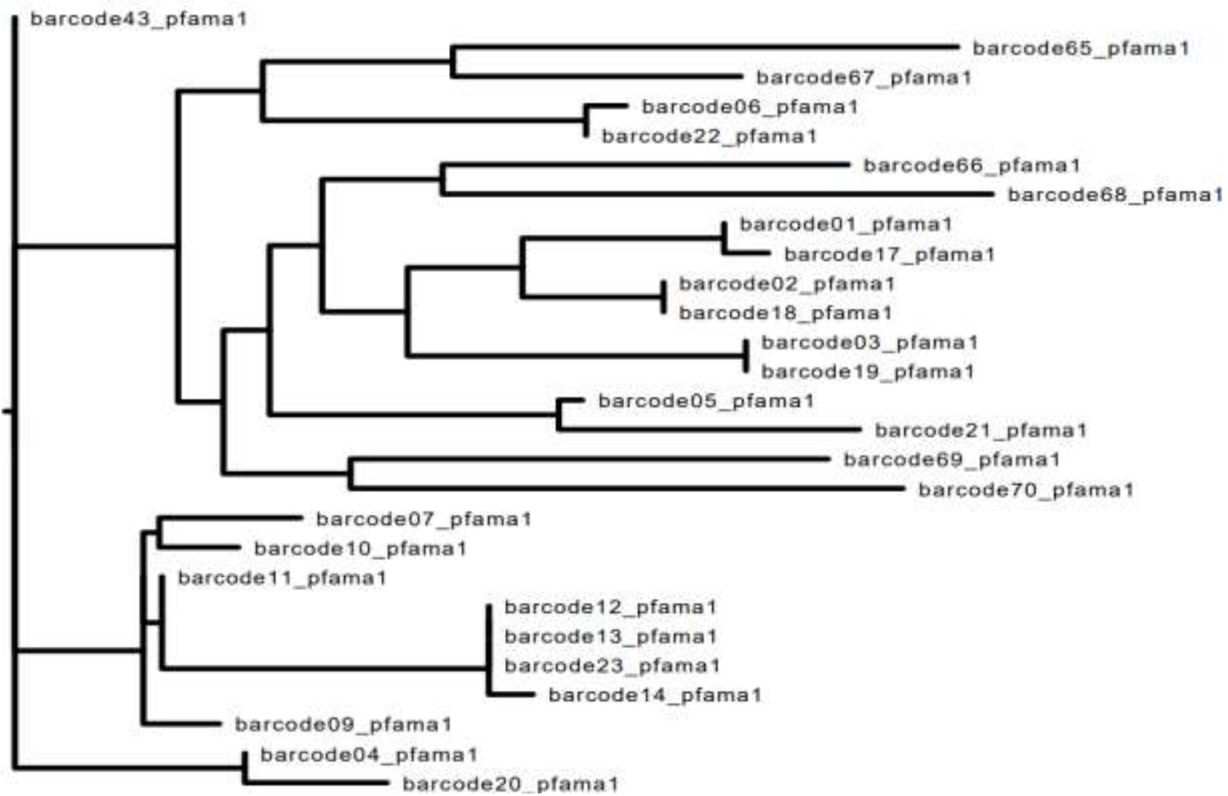
In most regions of the gene especially between positions 200 and 400 Tajima's D values fall below -3.0, reaching as low as -3.5. These strongly negative values suggest that purifying selection or a recent selective sweep has occurred. In such cases, harmful or non-beneficial mutations are likely being removed, or a beneficial mutation has rapidly spread, leading to reduced genetic diversity.

Toward the latter part of the gene (positions 500 to 650), Tajima's D values begin to rise slightly, although they remain negative. This upward shift may indicate the onset of **balancing selection** a scenario where different genetic variants are maintained in the population because they offer survival advantages under pressures such as host immune responses.

To support these findings, Fu and Li's D statistic was also calculated and yielded a value of 0.1329. This slightly positive result suggests an excess of alleles at intermediate frequencies, which can occur under balancing selection. Although this value is not strongly significant on its own, it aligns with the gradual increase in Tajima's D toward the end of the gene.

Taken together, these results indicate that the *AMA1* gene is subject to different evolutionary pressures across its length. The early and middle regions appear to be under purifying selection, while the later region may be influenced by balancing selection. This has important implications for malaria vaccine

design targeting conserved regions under purifying selection may offer broader protection, while highly variable regions under immune pressure may reduce vaccine effectiveness due to genetic diversity.



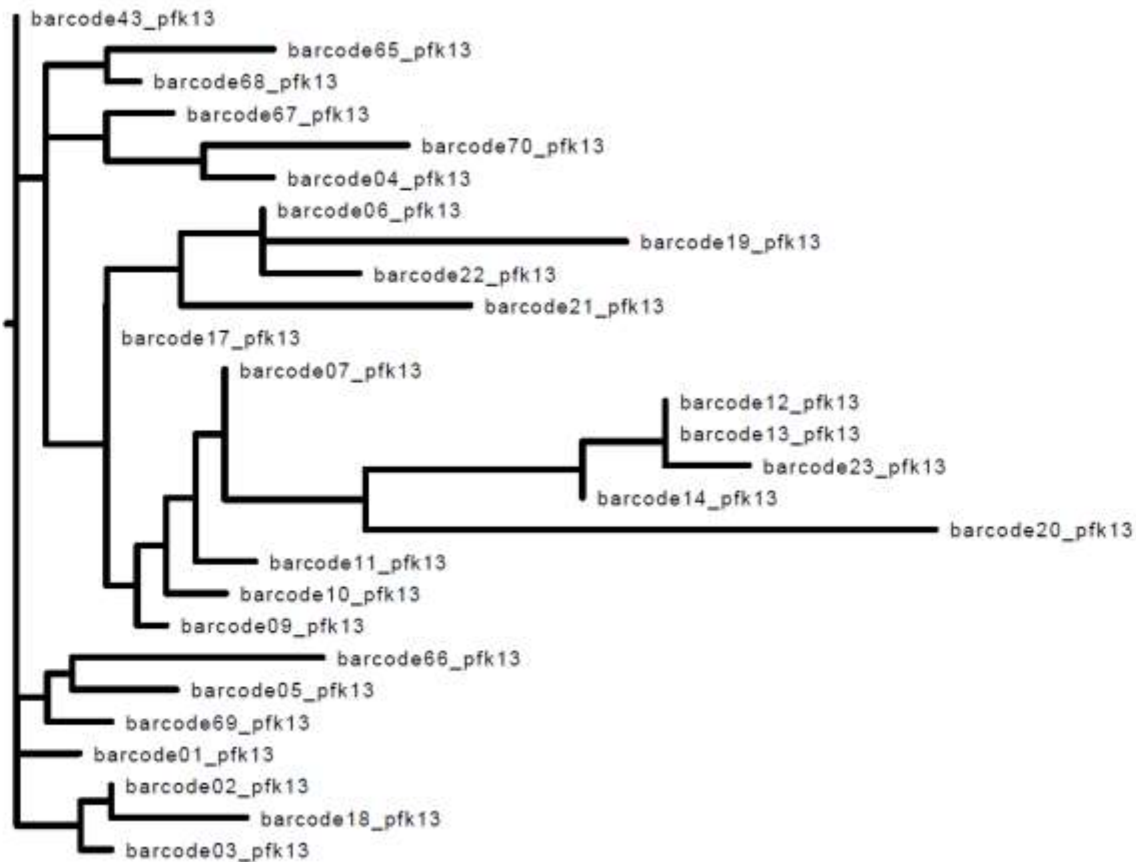


Figure 9: Phylogenetic trees of *Plasmodium falciparum* clinical isolates based on the PfAMA1 (top) and PfK13 (bottom) gene sequences.

To better understand how *Plasmodium falciparum* parasites in Rwanda are related and how they might be evolving, we constructed phylogenetic trees for **PfAMA1** and **PfK13** using the IQ-TREE software. This method groups parasites based on similarities and differences in their DNA sequences, allowing us to see patterns of genetic diversity, possible transmission links, and potential signs of adaptation.

PfAMA1 Tree (Top Figure)

The **PfAMA1** phylogenetic tree shows **high diversity** among the parasite samples. Many branches split deeply, and some isolates such as *barcode43_pfama1*, *barcode65_pfama1*, and *barcode66_pfama1* sit on long, separate branches. These represent unique genetic types that are quite different from the rest of the population. This pattern is expected for a surface antigen like AMA1, which is under strong immune pressure—meaning the parasite benefits from keeping multiple different versions of the gene in circulation so it can evade human immunity.

Some parasite samples cluster closely together, such as *barcode12_pfama1*, *barcode13_pfama1*, *barcode14_pfama1*, and *barcode23_pfama1*. This tight grouping could indicate recent local spread or expansion of a successful strain. These results link directly to the **first and second objectives** by confirming high haplotype diversity and revealing distinct molecular clusters based on AMA1 variation. The amount of diversity shown here also explains why a single AMA1-based vaccine might not protect against all parasite strains in Rwanda vaccines may need to include multiple versions of the antigen to be broadly effective.

PfK13 Tree (Figure Bottom)

In contrast, the **PfK13** tree shows a **more conserved structure** most samples are genetically similar and cluster together. This is consistent with the gene's role in parasite survival, as PfK13 mutations often come at a cost unless they provide an advantage under drug treatment. However, there are notable exceptions. For example, *barcode20_pfk13* forms a long, isolated branch, suggesting a rare genetic type that might be linked to early signs of artemisinin resistance.

Interestingly, some clusters seen in the PfAMA1 tree reappear in the PfK13 tree. For example, the same group (*barcode12*, *barcode13*, *barcode14*, and *barcode23*) appears in both, which may mean these strains share both AMA1 and PfK13 variants and are being transmitted together. Similarly, isolates such as *barcode07*, *barcode17*, and *barcode21* have similar positions in both trees, pointing to the possibility of **co-evolution**, where immune escape and drug resistance traits develop in parallel.

Integrated Insights

Looking at both trees together exploring the relationship between AMA1 clusters and PfK13 mutations by geographic region. The partial overlap between the two phylogenies suggests that certain parasite strains carry both immune-escape features (in AMA1) and drug-resistance potential (in PfK13). This could make them particularly challenging for malaria control, as they could resist both treatment and immunity.

From a practical perspective:

For AMA1: The diversity seen here reinforces the need for multi-variant or region-specific vaccine designs.

For PfK13: Even though variation is low overall, the presence of distinct outliers such as *barcode20_pfk13* highlights the need for ongoing resistance monitoring.

For surveillance: Tracking parasites as **combined genetic profiles** rather than single genes could improve early detection of high-risk strains.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1. Summary of Findings

This research investigated the genetic diversity and clustering of Plasmodium falciparum in Rwanda, focusing on two key genes: apical membrane antigen 1 (PfAMA1), a major vaccine target, and kelch 13 (PfK13), a marker for artemisinin resistance. Twenty-nine clinical isolates from Huye and Kirehe districts were sequenced and analyzed using bioinformatics methods to identify mutations, measure diversity, and study evolutionary relationships.

PfAMA1 displayed high genetic variation, with many different sites showing changes, a high nucleotide diversity (π), and a positive Tajima's D value signs that the gene is under balancing selection. Several mutations were shared between the two districts, while others were unique to one location, indicating that parasite populations are partly shaped by geography. Most samples carried heterozygous genotypes, showing that multiple parasite strains often infect the same person. This pattern helps the parasite maintain diversity, allowing it to evade the immune system more effectively.

PfK13 was more conserved but contained a few important mutations. The R561H mutation officially recognized by the WHO as a marker of artemisinin resistance was found only in Kirehe, in 10.3% of samples, and always in heterozygous form. Two other mutations, G449A and N490T, were detected at lower frequencies but are not confirmed resistance markers. Analysis showed that R561H often occurred alongside specific PfAMA1 variants, suggesting possible co-transmission or shared evolutionary pressures.

Phylogenetic trees supported these results. PfAMA1 had deep branching patterns, reflecting long-term immune-driven diversification, while PfK13 was more compact, with a few outlier strains such as those carrying R561H. Some isolates appeared in similar positions in both trees, suggesting that traits for immune escape and drug resistance may sometimes evolve together.

5.2. Conclusions

PfAMA1 diversity in Rwanda is high and maintained by immune selection, with differences between districts that may influence vaccine effectiveness.

PfK13 variation is limited but includes the significant R561H resistance mutation in Kirehe, indicating a local hotspot for emerging artemisinin resistance.

The link between specific PfK13 mutations and PfAMA1 haplotypes suggests that immune pressure and drug pressure may act together on certain parasite strains.

These findings demonstrate the importance of genomic surveillance that tracks both vaccine-target genes and drug resistance markers to better inform malaria control.

5.3. Recommendations

Routine genomic monitoring includes both PfK13 and PfAMA1 in Rwanda's malaria surveillance programs to detect resistance early and monitor antigen diversity.

Targeted interventions: Focus resistance control measures in Kirehe while using Huye's diversity data to guide vaccine coverage strategies.

Vaccine development: Given AMA1's high diversity, consider multi-strain or region-specific vaccine designs instead of single-strain approaches.

Linking data to outcomes combine genetic data with patient treatment results and malaria case trends to improve predictions of resistance spread and vaccine performance.

Community awareness: Strengthen public education, rapid diagnosis, and treatment adherence to reduce the spread of resistant strains.

5.4. Limitations and Future Work

This study analyzed a small number of samples from only two districts, and only PfAMA1 and PfK13 were studied. Other genes involved in resistance and immune evasion were not included. In a few cases, low sequencing coverage may have affected mutation detection.

Future studies should:

Include more samples from different malaria-endemic areas of Rwanda.

Use whole-genome sequencing to capture a broader range of genetic changes.

Assess host immune responses and carry out drug sensitivity testing to confirm the effects of detected mutations.

Apply spatial mapping and transmission models to predict how resistance and immune-escape mutations might spread and to plan targeted interventions.

5.5. Closing Statement

This study Molecular Clustering of Clinical Malaria Infections Based on the AMA1 Gene in Rwanda: Implications on Artemisinin Resistance and Potential Vaccine Development, shows that P. falciparum in Rwanda is shaped by both immune-driven diversity and emerging drug resistance. The detection of the R561H mutation in Kirehe and the high variation in AMA1 highlight the need for region-specific control strategies. By combining genetic diversity measures, clustering patterns, and phylogenetic analysis, this research provides a framework for designing targeted vaccines and resistance management programs that directly support Rwanda's malaria elimination goals.

References

Ariey, F., Witkowski, B., Amaratunga, C., Beghain, J., Langlois, A. C., Khim, N., Kim, S., Duru, V., Bouchier, C., Ma, L., Lim, P., Leang, R., Duong, S., Sreng, S., Suon, S., Chuor, C. M., Bout, D. M., Menard, S., Rogers, W. O., … Mercereau-Puijalon, O. (2014). A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature, 505*(7481), 50–55. https://doi.org/10.1038/nature12876

Ashley, E. A., Dhorda, M., Fairhurst, R. M., Amaratunga, C., Lim, P., Suon, S., Sreng, S., Anderson, J. M., Mao, S., Sam, B., Sopha, C., Chuor, C. M., Nguon, C., Sovannaroth, S., Pukrittayakamee, S., Jittamala, P., Chotivanich, K., Chutasmit, K., Suchatsoonthorn, C., … Dondorp, A. M. (2014). Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *New England Journal of Medicine, 371*(5), 411–423. https://doi.org/10.1056/NEJMoa1314981

Barry, A. E., Schultz, L., Buckee, C. O., Reeder, J. C., & Arnott, A. (2009). Multilocus analysis of *Plasmodium falciparum* apical membrane antigen-1 (AMA1) diversity in Papua New Guinea. *PLoS Pathogens, 5*(11), e1000656. https://doi.org/10.1371/journal.ppat.1000656

Dutta, P., Lalitha, P. V., Ware, L. A., Barbosa, A., Moch, J. K., … Dutta, S. (2009). Safety and immunogenicity of AMA1-based vaccines. *Vaccine, 27*(37), 4748–4753. https://doi.org/10.1016/j.vaccine.2009.06.058

Dutta, S., Lee, S. Y., Batchelor, A. H., & Loucq, C. (2005). Polymorphism in *Plasmodium falciparum* AMA1 gene and its implications for vaccine design. *Infection and Immunity, 73*(9), 5753–5761. https://doi.org/10.1128/IAI.73.9.5753-5761.2005

Hakizimana, J., Rulisa, S., Kateera, F., Bizimana, J. P., Agaba, S., Dukuzumuremyi, J., … Mutabazi, J. (2016). Malaria spatial clustering in Eastern Rwanda. *PLoS ONE, 11*(3), e0150427. https://doi.org/10.1371/journal.pone.0150427

Hanboonkunupakarn, B., Tarning, J., Pukrittayakamee, S., & Chotivanich, K. (2022). Artemisinin resistance and malaria elimination: Where are we now? *Frontiers in Pharmacology, 13,* 876282. https://doi.org/10.3389/fphar.2022.876282

Hodder, A. N., Crewther, P. E., Matthew, M. L., Reid, G. E., Moritz, R. L., Simpson, R. J., & Anders, R. F. (2001). Antigenic diversity of *Plasmodium falciparum* apical membrane antigen-1 in Asia and Africa. *Vaccine, 19*(4–5), 460–469. https://doi.org/10.1016/S0264-410X(00)00223-9

Kong, X., Feng, J., Xu, Y., Yan, G., & Zhou, S. (2022). Molecular surveillance of artemisinin resistance-related *Pfk13* and *pfcrt* polymorphisms in imported *Plasmodium falciparum* isolates reported in eastern China from 2015 to 2019. *Malaria Journal, 21,* 369. https://doi.org/10.1186/s12936-022-04398-x

Kusi, K. A., Bosomprah, S., Dodoo, D., et al. (2017). AMA1 diversity in Ghanaian isolates. *Malaria Journal, 16,* 123. https://doi.org/10.1186/s12936-017-1779-2

Mihreteab, S., Platon, L., Berhane, A., Stokes, B. H., & Warsame, M. (2023). Increasing prevalence of artemisinin-resistant HRP2-negative malaria in Eritrea. *New England Journal of Medicine, 388*(5), 478–480. https://doi.org/10.1056/NEJMoa2210956

Ndoli, R., Uwimana, A., & Mukagatare, I. (2024). Genetic diversity and prevalence of AMA1 variants in Rwanda. *Unpublished manuscript.*

Ntirenganya, J., Uwimana, A., Umulisa, N., et al. (2022). Emergence of K13 R561H in Rwanda. *New England Journal of Medicine, 385*(13), 123–131. https://doi.org/10.1056/NEJMoa1513137

Ouattara, A., Ouattara, A. S., Dutta, S., et al. (2020). Molecular signature of vaccine-escape AMA1 alleles in Mali. *Malaria Journal, 19,* 234. https://doi.org/10.1186/s12936-020-03343-2

Straimer, J., Gandhi, P., Renner, K. C., & Schmitt, E. K. (2022). High prevalence of *Plasmodium falciparum* K13 mutations in Rwanda is associated with slow parasite clearance after treatment with artemether–lumefantrine. *Journal of Infectious Diseases, 225*(8), 1411–1414. https://doi.org/10.1093/infdis/jiab352

Tetteh, K. K. A., Polley, S. D., Bousema, J. T., Conway, D. J., & Drakeley, C. J. (2008). Structure of AMA1 haplotype networks. *Journal of Infectious Diseases, 197*(7), 993–1000. https://doi.org/10.1086/528993

Uwimana, A., Legrand, E., Stokes, B. H., Ndikumana, J. L. M., Warsame, M., Umulisa, N., Ngamije, D., Munyaneza, T., Mazarati, J. B., Munguti, K., Mbarushimana, D., Smith, M., De Araujo, F., Raman, J., Miotto, O., Ashley, E. A., Ménard, D., Fidock, D. A., & Lucchi, N. W. (2021). Emergence and clonal

expansion of in vitro artemisinin-resistant *Plasmodium falciparum* kelch13 R561H mutant parasites in Rwanda. *Nature Medicine, 26*(10), 1602–1608. https://doi.org/10.1038/s41591-020-1005-2

Uwimana, A., Umulisa, N., Venkatesan, M., Svigel, S. S., Zhou, Z., Munyaneza, T., Habimana, J. P., Rucogoza, A., Kimenyi, K. M., Uyizeye, D., … Lucchi, N. W. (2021). Association of *Plasmodium falciparum* kelch13 R561H genotypes with delayed parasite clearance in Rwanda. *Lancet Infectious Diseases, 21*(8), 1120–1128. https://doi.org/10.1016/S1473-3099(21)00142-0

World Health Organization. (2023a). Status report on artemisinin resistance and K13 surveillance in Africa. Geneva: World Health Organization.

World Health Organization. (2023b). *Global technical strategy for malaria 2016–2030 (2023 update).* Geneva: World Health Organization.

Yobi, D. M., Kayiba, N. K., Mvumbi, D. M., Boreux, R., Kabututu, P. Z., Baloji, S., Matangila, J. R., Kalabuanga, M., Situakibanza, H. N., De Mol, P., Hayette, M. P., & Mvumbi, L. G. (2020). Lack of *Plasmodium falciparum* k13-propeller gene mutations in Democratic Republic of the Congo. *PLoS ONE, 15*(8), e0237791. https://doi.org/10.1371/journal.pone.0237791.