CHAPTER ONE

1.0 INTRODUCTION

Malaria is a parasitic disease that is responsible for morbidity in approximately 214 million people across the world and an estimated 438, 000 deaths, most of which occur in sub-Saharan Africa (sSA) (WHO, 2015). There exists a high level of malaria burden caused by *Plasmodium falciparum* infection in Nigeria with over 170 million people at risk (FMoH, 2013). Malaria prevalence in the country is highest in the South-West geopolitical zone (NMIS, 2010), despite more than a decade of vector control with Insecticide-Treated Nets/Long-Lasting Insecticidal Nets (ITN/LLINs), Indoor Residual Spraying (IRS), larval control and targeting of parasites with Intermittent Preventive Treatment (IPT) and Artemisinin-based Combination Therapy (ACT).

Contributing to the burden of the disease is drug resistance that has crippled most antimalarial drugs (Chukwuocha *et al*., 2013; Olasehinde *et al*., 2014). In order to track the efficacy of existing antimalarial drugs, in vivo evaluation of drugs is carried out over a follow-up period (Happi *et al*., 2004). Molecular genotyping of pre-treatment (baseline) and recurrent infections using established antigenic markers (Merozoite-Surface Proteins, MSP-1 and MSP2) enables the categorization of recurrent parasites as recrudescent (true failure) or new infection (successful treatment) either from pre-existing infection or a new infection from an infected mosquito bite since the probability of a patient to be newly infected with a parasite possessing an identical genotype to the former infection is low. Therefore, comparing the genotypes of established antigenic markers at baseline and at the time of parasite recurrence is expected to discriminate between recrudescent and new infections (Snounou *et al*., 1998).

However, the discriminatory power of these antigenic markers is dependent on the frequency and polymorphism of each allele of the MSP genotypes within the parasite population (Snounou *et al*., 1998). An infection appearing during follow-up after the parasites have been cleared can be incorrectly classified as recrudescent if there is low clonal distribution of the parasites because a new infection may likely share the same genotype as the baseline infection. This may cause an over-estimation of treatment failures and unnecessary treatment policy changes.

Genetic diversity studies of *P. falciparum* provide information needed to identify targets of immunity as potential vaccine candidates, and to monitor the emergence and evolution of drug resistance. Advances in parasite population genetics approaches have increased the efficiency of detecting pressure arising from host-parasite interactions (Amambua-Ngwa *et al*., 2012; Mobegi *et al*., 2014). Such interfaces between the parasite and human host's immune system leave selections in the genes responsible for drug resistance and/or evasion of host's immunity (Fumagalli *et al*., 2009).

A number of such genes encoding drug or immune targets have been linked with natural selection in *P. falciparum* identified as signatures of directional or balancing selections (Escalante *et al*., 1998; Volkman *et al*., 2007; Mu *et al*., 2007; Weedal *et al*., 2007). Directional forces cause adaptively important genetic variants of the parasite to increase in frequency leading to high fixation rates, the appearance of a selective sweep and reduced variability in the region flanking the selected loci. Output from proteomic (Florens *et al*., 2002; Bozdech *et al*., 2003; Le Roch *et al.*, 2003; Crosnier *et al*., 2011), genomic (Mu *et al*., 2010; Le Roch *et al.*, 2012) and population genetic studies (Conway *et al*., 2000; Mu *et al.*, 2007; Polley *et al*., 2007; Tetteh *et*

al., 2009) of *P. falciparum* from different endemic regions as well as previous analyses on target microsatellites and single nucleotide polymorphisms (SNPs) in malaria endemic regions have shown strong selective forces operating around several drug resistance genes (Anderson *et al*., 2000; Volkman *et al*., 2007; Mu *et al*., 2010).

However, balancing selection brings the parasite's favoured alleles to an intermediate equilibrium, where they are maintained as genetic polymorphisms that potentially increase the variability in the genomic region flanking the selected loci. Studies on the statistical distribution of DNA polymorphism within and among populations have indicated strong signatures of balancing selection on particular vaccine candidate antigens including MSP-1 (Conway *et al*., 2000), MSP2 (Ferreira and Hartl, 2007), erythrocyte binding antigen 175 (EBA-175) (Verra *et al*., 2006) and MSP3 (Polley *et al*., 2007). Thus systematic scan for signatures of balancing selection has become imperative to prospect for new vaccine developments based on multi-allelic antigen devising.

Next-generation genomic technologies have proved useful in individual sequencing of *P. falciparum* in order to identify genes with polymorphic site frequency spectra consistent with selection (Mu *et al*., 2007). However, such technologies are still infeasible in resource-poor settings where malaria is endemic and large data sets are required. In such regions high-throughput sequencing of samples pooled from different individual isolates (pool-seq) has been used to detect Single Nucleotide Polymorphisms (SNPs) with a variance comparable with individual sequencing (Ferretti *et al*., 2013). However, this strategy has not been effectively

adopted in *P. falciparum* genomics. It is important to understand the genome-wide patterns of parasite selection in order to identify genes that are most variable and potentially under strong directional and balancing selection from antimalarial drug use and host immunity.

1.1 STATEMENT OF THE PROBLEM

Malaria is a major health concern in Nigeria accounting for 25% of infant mortality, 30% of childhood deaths and 11% maternal mortality (FMoH, 2013). Compounding the disease burden is high level of genetic variability of the parasites which makes them evasive to control efforts. Blood-stage *P. falciparum* infections usually harbour multiple haploid clones which may vary in immunogenicity and susceptibility to drugs. Clonal diversity of malarial infections not only presents a major barrier to vaccine development (Kilama and Ntoumi, 2009), it also brings about challenges in drug resistance tracking (Takala and Plowe, 2009).

Molecular genotyping of pre- and post-treatment infections using MSP1 and MSP2 enables categorization of infections as recrudescence or re-infection. The accuracy of this approach is dependent on the allelic variants and clonal diversity of the parasites within a population. Inadequate information on the clonal diversity of *P. falciparum* within a population may cause a poor estimation of treatment failures and unnecessary drug policy changes.

Emergence and spread of novel *P. falciparum* phenotypes such as drug resistance or antigenic variants within and between populations is dependent on the extent of genetic diversity and structure of the parasites. There is little information on parasite population trends and genetics needed to help guide control programmes in Nigeria, hence the need to understand current

structure and the extent of genetic diversity across different populations. There is also paucity of information on candidate antigens that can form the basis for vaccine development against Nigerian populations of *P. falciparum*. It is important to identify those genes that are most variable and potentially under strong directional and balancing selection from antimalarial drug use and host immunity. This will allow effective translation of *P. falciparum* genomics into health interventions useful for the control of malaria.

1.2 AIM OF THE STUDY

The aim of the study is to understand the population structure and genomics of *P. falciparum* isolates in Lagos and Ekiti States, southwestern Nigeria.

1.3 OBJECTIVES OF THE STUDY

The objectives are to:

1. Determine the allelic variants, clonal diversity, multiplicity of infections and genetic differentiation of *P. falciparum* isolates in southwestern Nigeria.
2. Determine the population structure of *P. falciparum* isolates
3. Establish the extent of directional selection on genes associated with antimalarial drug resistance
4. Identify target genes of immunity under balancing selection as candidates of malaria vaccines

1.4 SIGNIFICANCE OF THE STUDY

This investigation will provide an insight into the genetic diversity and structure of *P. falciparum*. Genome-wide scan of the parasite isolates will provide information on the impact of selection pressure on common antimalarial target genes and also predict antigens that can serve

as potential vaccine candidates. This will allow the design of evidence-based malaria intervention strategies.

1.5 RESEARCH QUESTIONS

1. What is the extent of polymorphism in the MSP-1 and MSP2 antigenic markers?
2. Are the *P. falciparum* populations in southwestern Nigeria structured?
3. What is the extent of multiplicity of infections with *P. falciparum*?
4. Is there any evidence of recent selective sweeps around genes associated with antimalarial drug resistance?
5. How strong are the selective forces, if present?
6. Which genes in the Nigerian isolates of *P. falciparum* are under the effect of balancing selection?

1.6 DEFINITION OF TERMS

Balancing selection: selection in which polymorphic alleles are actively maintained at equilibrium in the gene pool of a population at frequencies above that of gene mutation.

Clonal diversity: The existence of high frequency of monoclonal infections

Directional selection: A type of natural selection in which an extreme phenotype of the parasite (e.g drug resistance) is favoured causing the allele frequency to shift over time in the direction of that phenotype.

Genome: An organism's complete set of DNA, including all of its genes.

Integrated haplotype score: A measure of the amount of extended haplotype homozygosity (EHH) at a given SNP along the ancestral allele relative to the derived allele.

Natural selection: The differential survival and reproduction of individuals due to differences in phenotype

Nonsynonymous Polymorphism: Change in the amino acid composition when the reference allele is substituted with the variant alleles

Selective sweep - reduction or elimination of variation among the nucleotides in neighboring DNA of a mutation as a result of recent and strong positive natural selection.

Synonymous Polymorphism: No change in the amino acid composition when the reference allele is substituted with the variant allele.

Tajima's D: A statistical test used to distinguish between a DNA sequence evolving randomly and one evolving under a non-random process.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 THE BURDEN OF MALARIA

Malaria causes morbidity among approximately 214 million people across the globe and an estimated 438, 000 deaths. Approximately half of the world's population is at risk of malaria. Ninety percent of malaria cases and deaths occur in sub-Saharan Africa especially among children where a child dies every minute from malaria (WHO, 2015). Ninety-seven countries have ongoing malaria transmission (Figure 2.1) covering Africa, Asia, Latin America, and to a lesser extent the Middle East and parts of Europe (WHO, 2014). Nigeria contributes about 32% of the global burden of the disease (WHO, 2015). The prevalence of the disease varies by geopolitical zone with the highest prevalence found in the South-West (50%), North-Central (49%) and North-West (48%), while the lowest prevalence zones are South-East (28%), North-East (31%) and South-South (32%) (NMIS, 2010).

2.2 HUMAN MALARIA PARASITE SPECIES AND DISTRIBUTION

2.2.1 *Plasmodium vivax*

Plasmodium vivax is the most frequent and widely distributed cause of recurring (benign tertian) malaria. It has been estimated that 2.5 billion people are at risk of infection with this parasite (Gething *et al*., 2012). Although the Americas contribute to the global area at risk, high endemic areas are generally sparsely populated and the region contributes only 6% to the total population at risk. In Africa, the widespread lack of the Duffy antigen in the population has ensured that stable transmission is constrained to Madagascar and parts of the Horn of Africa. It contributes 3.5% of global population at risk. Central Asia is responsible for 82% of global population at risk with high endemic areas coinciding with dense populations particularly in India and Myanmar.

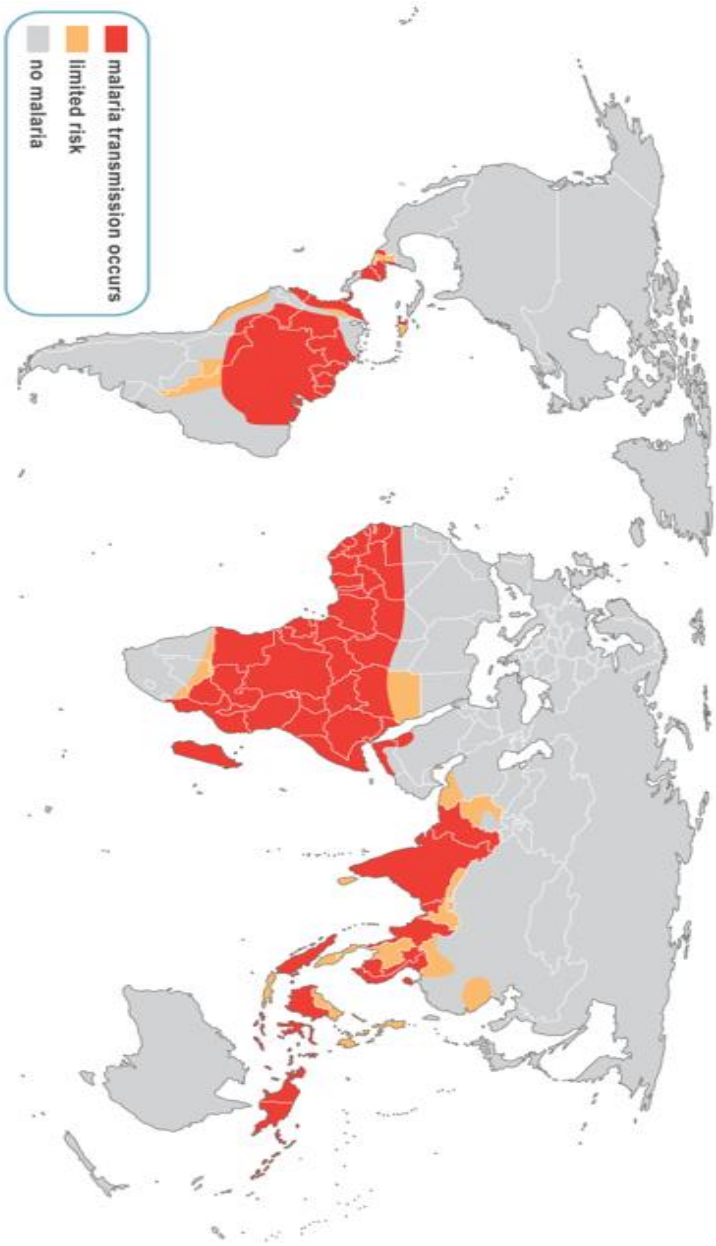

Figure 2.1: Global distribution of malaria (WHO, 2014)

South East Asia has areas of high endemicity in Indonesia and Papua New Guinea and overall contributes 9% of global population at risk (Battle *et al*., 2012). *P. vivax* is carried by at least 71 mosquito species a lot of which prefer to bite outdoors or during the daytime, hampering the effectiveness of indoor insecticide and bed nets.

Plasmodium vivax is divided into two subtypes, a dominant form, VK210 and a variant form, VK247. This division is dependent on the amino acid composition of the circumsporozoite (CS) protein. A strain of *P. vivax* containing a variant repeat in its CS protein was first isolated in Thailand (Tong-Soo *et al*., 2010). The CS repeat of this variant strain (Thai VK247) differs at 6/9 amino acids within the repeat sequence found in all previously described *P. vivax* CS protein. Following this discovery, several studies have been conducted to evaluate the global distribution of variant VK247; it was detected in indigenous populations of China (Han *et al*., 1999), Brazil (Branquinho *et al*., 1993), Mexico (Kain *et al.*, 1992), Peru (Need *et al*., 1993), and Papua New Guinea (Kain *et al*., 1993).

The drug susceptibility of the VK247 subtype of *P. vivax* is slightly different from VK210 (Kain *et al*., 1993), as well as that *Anopheles albimanus* and *An. pseudopunctipennis* differ in their susceptibilities to *P. vivax* circumsporozoite phenotypes. *An. albimanus* is more susceptible to the VK210 subtype, whereas *An. pseudopunctipennis* is more susceptible to the VK247 subtype (Gonzalez-Ceron *et al*., 1999).

2.2.2 *Plasmodium malariae*

Plasmodium malariae is one of the least studied malaria species that infects humans, in part because of its low prevalence and milder clinical manifestations compared to the other species

(Bruce *et al*., 2006). It is widespread throughout sub-Saharan Africa, much of South-East Asia, Indonesia, on many of the Islands of the Western Pacific and in areas of the Amazon Basin of South America (Westling *et al.*, 1997). In endemic regions, prevalence ranges from less than 4% to more than 20% (Bruce *et al*., 2006) but there is evidence that *P. malariae* infections are vastly under-reported (Mohapatra *et al*., 2008).

2.2.3 *Plasmodium ovale*

Plasmodium ovale has been shown by genetic methods to consist of two subspecies, *P. ovale curtisi* (classic type) and *P. ovale wallikeri* (variant type) (Sutherland *et al*., 2010), named after malaria researchers Christopher Curtis (1939-2008) and David Walliker (1940-2007). These two non-recombining, genetically distinct species co-exist, being sympatric in Africa and Asia. Splitting of the 2 lineages is estimated to have occurred between 1.0 and 3.5 million years ago in hominid hosts (Sutherland *et al*., 2010). The two subspecies appear to differ in their biology with *P. ovale wallikeri* having a shorter latency period than *P. ovale curtisi.*

2.2.4 *Plasmodium knowlesi*

Plasmodium knowlesi infection is normally considered a parasite of long-tailed (*Macaca fascicularis*) and pig-tailed *(Macaca nemestrina*) macaques but humans who work at the forest fringe or enter the rainforest to work are at risk of infection (Ng *et al*., 2008; Vythiligam *et al*., 2008). With the increasing popularity of deforestation and development efforts in South East Asia, many macaques are now coming in close and direct contact with humans (Vythiligam *et al*., 2008). Hence more people who live in the semi-urban areas are found to be infected with *knowlesi* malaria. This parasite is mostly found in South East Asian countries particularly in

Borneo, Cambodia, Malaysia, Myanmar, Philippines, Singapore, Thailand and neighbouring countries and it appears to occur in regions that are reportedly free of the other four types of human malaria (Jeslyn *et al*., 2011; Jongwutiwes *et al*., 2011). Infective mosquitoes are restricted to the forest areas. Non-infective mosquitoes are typically found in the urban areas but transmission may occur due to the abundance of mosquitoes in this region particularly Malaysia (Singh *et al*., 2004).

There have been reports of infection on the Thai-Burmese border (Jonwutiwes *et al*., 2011). *P. knowlesi* is the most common cause of malaria in childhood in the Kudat district of Sabah, Malaysia (Barber *et al*., 2011). One fifth of the cases of malaria diagnosed in Sarawak, Malaysian Borneo are due to *P. knowlesi* (Singh *et al*., 2004). *Plasmodium knowlesi* has not been found in Africa. This may be because there are neither long-tailed nor pig-tailed macaques (the reservoir hosts of *P. knowlesi*) in Africa, and many West Africans lack the Duffy antigen, a protein on the surface of the red blood cell that the parasite uses to invade.

2.2.5 *Plasmodium falciparum*

Plasmodium falciparum parasite causes the most severe form of malaria, and is found in many tropical and subtropical regions of the world. It is responsible for 85% of the malaria cases and is the deadliest parasitic disease killing over one million people each year. Ninety percent of the deaths occur in sub-Saharan Africa among under five-year-old children. In addition to Africa, *falciparum* malaria occurs in South and Southeast Asia, Central and South America, the Caribbean and the Middle East (WHO, 2014).

2.3 BIOLOGY AND TRANSMISSION OF FALCIPARUM MALARIA

The malaria parasite exhibits a complex life cycle involving an insect vector (mosquito) and a vertebrate host (human). Infection is initiated when sporozoites are injected with the saliva of a feeding mosquito. Sporozoites are carried by the circulatory system to the liver and invade hepatocytes. The intracellular parasite undergoes an asexual replication known as exo-erythrocytic schizogony within the hepatocyte. Exo-erythrocytic schizogony culminates in the production of merozoites which are released into the bloodstream. A proportion of the liver-stage parasites from *P. vivax* and *P. ovale* go through a dormant period instead of immediately undergoing asexual replication. These hypnozoites will reactivate several weeks to months (or years) after the primary infection and are responsible for relapses. Merozoites invade erythrocytes and undergo a trophic period in which the parasite enlarges. The early trophozoite is often referred to as ring form because of its morphology.

Trophozoite enlargement is accompanied by an active metabolism including the ingestion of host cytoplasm and the proteolysis of hemoglobin into amino acids. The end of the trophic period is manifested by multiple rounds of nuclear division without cytokinesis resulting in a schizont. Merozoites emerge from the mature schizont, also called a segmenter, and the merozoites are released following rupture of the infected erythrocyte. Invasion of erythrocytes reinitiates another round of the blood-stage replicative cycle (Figure 2.2). The blood stage is responsible for the pathology associated with malaria.

The intermittent fever paroxyms are due to the synchronous lysis of the infected erythrocytes. *P. malariae* exhibits a 72 hour periodicity, whereas the other three species exhibit 48 hour cycles.

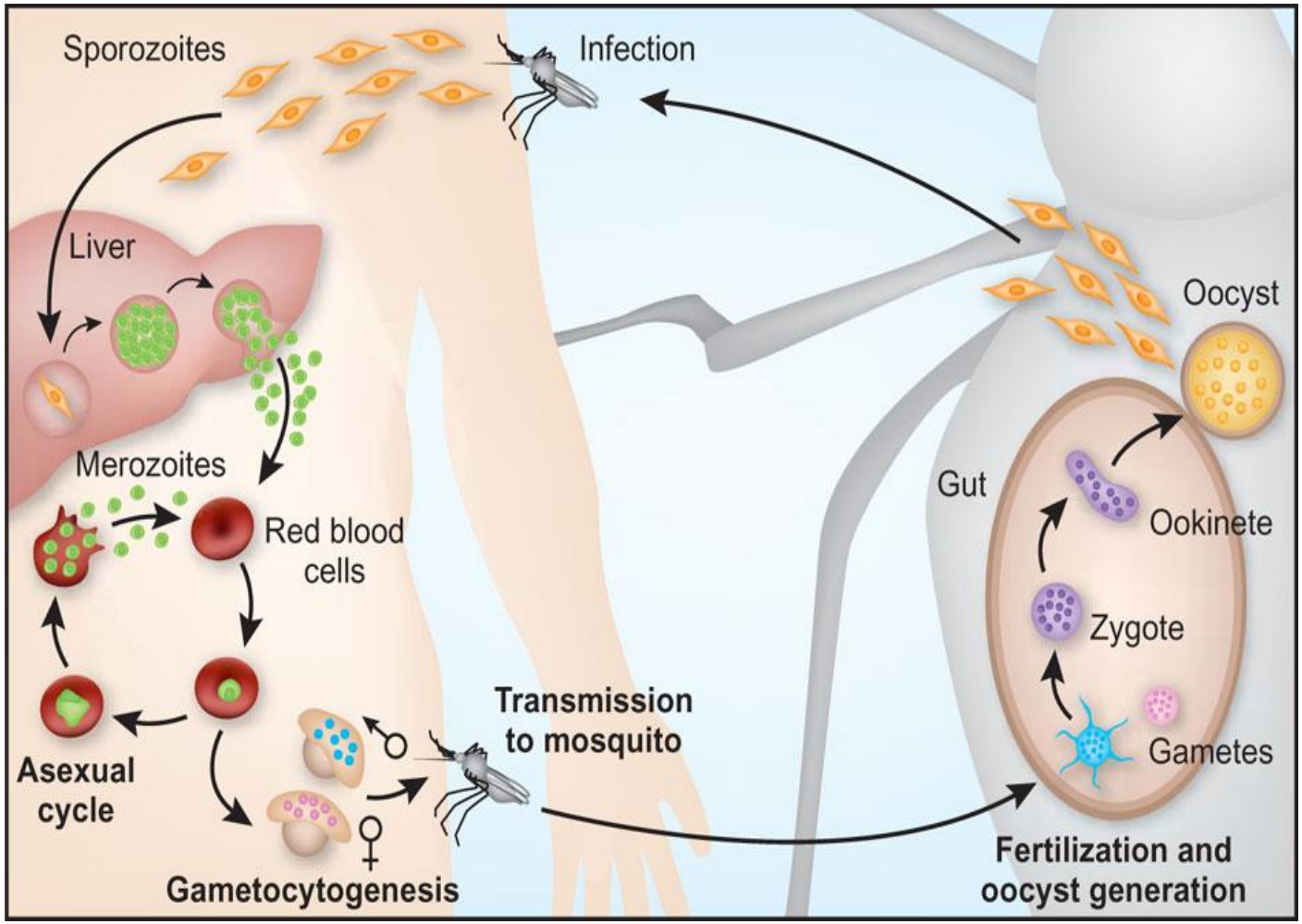

Figure 2.2 Life cycle of *Plasmodium falciparum*

However, *P. falciparum* often exhibits a continuous fever rather than the periodic paroxyms. *P. falciparum* also is responsible for more morbidity and mortality than the other species. This increased virulence is due in part to the higher levels of parasitaemia associated with *P. falciparum* infections. In addition, more complications are associated with *P. falciparum* because of the sequestration of the trophozoite- and schizont-infected erythrocytes in the deep tissues. As an alternative to the asexual replicative cycle, the parasite can differentiate into sexual forms known as macro- or microgametocytes. The gametocytes are large parasites which fill up the erythrocyte, but only contain one nucleus. Ingestion of gametocytes by the mosquito vector induces gametogenesis and escape from the host erythrocyte. Factors which participate in the induction of gametogenesis include: a drop in temperature, an increase in carbon dioxide, and mosquito metabolites.

Microgametes, formed by a process known as ex-flagellation, are flagellated forms which will fertilize the macrogamete leading to a zygote. The zygote develops into a motile ookinete which penetrates the gut epithelial cells and develops into an oocyst. The oocyst undergoes multiple rounds of asexual replication resulting in the production of sporozoites. Rupture of the mature oocyst releases the sporozoites into the haemocoel (i.e., body cavity) of the mosquito. The sporozoites migrate to and invade the salivary glands, thus completing the life cycle. Malaria parasites undergo three distinct asexual replicative stages (exo-erythrocytic schizogony, blood stage schizogony, and sporogony) resulting in the production of invasive forms (merozoites and sporozoites).

2.4 MALARIA DIAGNOSIS

The mainstay of malaria diagnosis has been the microscopic examination of blood, utilizing blood films (Krafts *et al*., 2011). Although blood is the sample most frequently used to make a diagnosis, both saliva and urine have been investigated as alternative, less invasive specimens (Sutherland and Hallett, 2009). More recently, modern techniques utilizing antigen tests or polymerase chain reaction have been discovered, though these are not widely implemented in malaria endemic regions (Ling *et al*., 1986; Mens *et al.*, 2006). Areas that cannot afford laboratory diagnostic tests often use only a history of subjective fever as the indication to treat for malaria.

2.4.1 Microscopy

The most economic, preferred, and reliable diagnosis of malaria is microscopic examination of blood films because each of the four major parasite species has distinguishing characteristics. Two sorts of blood film are traditionally used. Thin films are similar to usual blood films and allow species identification because the parasite's appearance is best preserved in this preparation. Thick films allow the microscopist to screen a larger volume of blood and are about eleven times more sensitive than the thin film, so picking up low levels of infection is easier on the thick film, but the appearance of the parasite is much more distorted and therefore distinguishing between the different species can be much more difficult. With the pros and cons of both thick and thin smears taken into consideration, it is imperative to utilize both smears while attempting to make a definitive diagnosis (Warhurst and Williams, 1996). From the thick film, an experienced microscopist can detect parasite levels (or parasitaemia) as few as 5 parasites/µL blood. Diagnosis of species can be difficult because the early trophozoites (ring

form) of all four species look identical and it is never possible to diagnose species on the basis of a single ring form; species identification is always based on several trophozoites. *P. malariae* and *P. knowlesi* (which is the most common cause of malaria in South-East Asia) look very similar under the microscope. However, *P. knowlesi* parasitaemia increases very fast and causes more severe disease than *P. malariae*.

2.4.2 Dipstick approach

For areas where microscopy is not available, or where laboratory staff are not experienced at malaria diagnosis, there are commercial antigen detection tests that require only a drop of blood (Pattanasin *et al*., 2003). Immuno-chromatographic tests (also called: Malaria Rapid Diagnostic Tests, Antigen-Capture Assay or Dipsticks) have been developed, distributed and field-tested. These tests use finger-stick or venous blood, the completed test takes a total of 15–20 minutes, and the results are read visually as the presence or absence of colored stripes on the dipstick, so they are suitable for use in the field. The threshold of detection by these rapid diagnostic tests is in the range of 100parasites/µl of blood (commercial kits can range from about 0.002% to 0.1% parasitemia) compared to 5parasites/µl by thick film microscopy. One disadvantage is that dipstick tests are qualitative but not quantitative.

The first rapid diagnostic tests relied on *P. falciparum* glutamate dehydrogenase (PGluDH) as antigen (Ling *et al*., 1986). PGluDH was soon replaced by *P. falciparum* lactate dehydrogenase (pLDH), a 33 kDa oxidoreductase. It is the last enzyme of the glycolytic pathway, essential for ATP generation and one of the most abundant enzymes expressed by *P. falciparum*. pLDH does not persist in the blood but clears about the same time as the parasites following successful

treatment. The lack of antigen persistence after treatment makes the pLDH test useful in predicting treatment failure. In this respect, pLDH is similar to pGluDH. Depending on which monoclonal antibodies are used, this type of assay can distinguish between all five different species of human malaria parasites, because of antigenic differences between their pLDH isoenzymes (Murray and Bennett, 2009).

2.4.3 Polymerase Chain Reaction (PCR) method

Molecular methods are available in some clinical laboratories and rapid real-time assays (Mens *et al*., 2006) are being developed with the hope of being able to deploy them in endemic areas. PCR (and other molecular methods) is more accurate than microscopy. However, it is expensive, and requires a specialised laboratory. Moreover, levels of parasitaemia are not necessarily correlative with the progression of disease, particularly when the parasite is able to adhere to blood vessel walls. Therefore more sensitive, low-tech diagnosis tools need to be developed in order to detect low levels of parasitaemia in the field (Redd *et al*., 1992). Another approach is to detect the iron crystal by-product of haemoglobin that is found in malaria parasites feasting on red blood cells, but not found in normal blood cells. It can be faster, simpler and precise than any other method.

2.4.4 Quantitative Buffy Coat

Quantitative buffy coat (QBC) is a laboratory test to detect infection with malaria or other blood parasites. The blood is taken in a QBC capillary tube which is coated with acridine orange (a fluorescent dye) and centrifuged; the fluorescing parasites can then be observed under ultraviolet light at the interface between red blood cells and buffy coat. This test is more sensitive than the

conventional thick smear; however it is unreliable for the differential diagnosis of species of parasite (Adeoye and Nga, 2007). In cases of extremely low white blood cell count, it may be difficult to perform a manual differential of the various types of white cells, and it may be virtually impossible to obtain an automated differential. In such cases a buffy coat may be obtained, from which a blood smear is made. This smear contains a much higher number of white blood cells than whole blood.

2.4.5 Loop-mediated isothermal amplification

Loop-mediated isothermal amplification (LAMP) is a nucleic acid amplification method that relies on autocycling strand-displacement DNA synthesis performed with Bst DNA polymerase. The amplification products are stem-loop DNA structures with several inverted repeats of the target and structures with multiple loops. The pricncipal merit of this method is that no denaturation of the DNA template is required (Nagamine *et al*., 2001), and thus, the LAMP reaction can be conducted in the field under isothermal conditions (ranging from 60 to 65^0C). LAMP requires only one enzyme and four primers that recognize six distinct target regions. The method produces a large amount of amplified product, resulting in easier detection, such as detection by visual judgement of the turbidity or fluorescence of the reaction mixture (Mori *et al.*, 2001).

Several investigators have reported on LAMP methods for the rapid identification of *Plasmodium*, *Trypanosoma* and *Babesia* (Ikadai *et al*., 2004; Poon *et al*., 2006; Thekisoe *et al*., 2005).

2.5 MALARIA CONTROL STRATEGIES

Several control measures are currently used against the human malaria parasites and their mosquito vectors. Such measures include: controlling the insect vector; early diagnosis and treatment of malaria.

2.5.1 Control of insect vectors

Measures to control the insect vector, reducing the numbers of mosquitoes and hence reducing transmission of the malaria parasite, include insecticide spraying, environmental management and biological control. Environmental management measures, such as covering wells and filling in ditches, and keeping irrigation channels fast flowing, help reduce the number of mosquito breeding sites.

2.5.1.1 Long-lasting Insectide-treated Nets

Long-lasting Insecticide-Treated Nets (LLINs) are an important public health strategy for malaria prevention adopted by most countries with endemic malaria. In addition to serving as physical barriers between mosquito vectors and individual users, toxicity and repellency induced by the pyrethroid insecticide-impregnated in LLINs can have important community-wide effects on vector density (Hawley *et al*., 2003), and LLINs have been shown to reduce the burden of malaria, especially among children under five years and pregnant women (Gamble *et al*., 2007) who are most vulnerable to malaria. LLINs are also one of the most cost-effective interventions, particularly in areas of high-malaria transmission (Wiseman *et al.*, 2003).

2.5.1.2 Larviciding

This is an ecologically safe preventive method used to interrupt the development of larvae or pupa into adult mosquitoes. Larvicing program emphasizes targeting mosquitoes in their habitats

where they reproduce. This gives the advantage of controlling the mosquitoes before mature to infective adults. Larviciding measures include the use of species of larvivorous fish and biolarvicides such as *Bacillus thuringiensis israelensis* and *Bacillus sphaericus* (Kumar *et al*., 1998).

2.5.1.3 Indoor Residual Spraying

Indoor residual spraying or IRS is the process of spraying the inside of dwellings with an insecticide to kill mosquitoes that spread malaria. A dilute solution of insecticide is sprayed on the inside walls of certain types of dwellings—those with walls made from porous materials such as mud or wood but not plaster as in city dwellings. Mosquitoes are killed or repelled by the spray, preventing the transmission of the disease. There are 12 insecticides recommended by WHO for IRS (WHO, 2006), belonging to four chemical groups (one organochlorine, six pyrethroids, three organophosphates and two carbamates). The choice of insecticide is informed by insecticide susceptibility and vector behavior, safety for humans and the environment and efficacy and cost-effectiveness (WHO, 2006). However, the potential threat of resistance to public health insecticides is significant.

2.5.2 Chemotherapy

Making early diagnosis and giving prompt treatment not only serves to cure the disease but to also reduce the spread of drug resistance. In low transmission areas, prompt treatment can interrupt malaria transmission (Chuma *et al*., 2010). Antimalarial drugs are used for the treatment and prevention of malaria infection. Most antimalarial drugs target the erythrocytic stage of malaria infection, which is the phase of infection that causes symptomatic illness. Treatment of the acute blood stage infection is necessary for disease caused by all malaria species. In addition, for infection due to *P. ovale* or *P. vivax,* terminal prophylaxis is required

with a drug active against hypnozoites (which can remain dormant in the liver for months, and occasionally years, after the initial infection).

2.5.2.1 Quinoline derivatives

Quinoline derivatives include chloroquine, amodiaquine, quinine, quinidine, mefloquine, primaquine, lumefantrine, and halofantrine. These drugs have activity against the erythrocytic stage of infection; primaquine also kills intrahepatic forms and gametocytes. The drugs act by accumulating in the parasite food vacuole and forming a complex with haeme that prevents crystallization in the plasmodium food vacuole. Haeme polymerase activity is inhibited, resulting in accumulation of cytotoxic-free haeme. Chloroquine was the first drug produced on a large scale for treatment and prevention of malaria infection. Chloroquine was, until recently, the most widely used anti-malarial. It was the original prototype from which most methods of treatment are derived. It is also the least expensive, best tested and safest of all available drugs. The emergence of drug-resistant parasitic strains is rapidly decreasing its effectiveness. Popular drugs based on chloroquine phosphate (also called nivaquine) are Chloroquine FNA, Resochin and Dawaquin. Chloroquine is a 4-aminoquinoline compound with a complicated and still unclear mechanism of action. It is believed to reach high concentrations in the vacuoles of the parasite, which, due to its alkaline nature, raises the internal pH. It controls the conversion of toxic haeme to haemozoin by inhibiting the bio-crystallization of haemozoin, thus poisoning the parasite through excess levels of toxicity. Other potential mechanisms through which it may act include interfering with the biosynthesis of parasitic nucleic acids and the formation of a chloroquine-haem or chloroquine-DNA complex. The most significant level of activity found is against all forms of the schizonts (with the obvious exception of chloroquine-resistant *P. falciparum* and *P.*

vivax strains) and the gametocytes of P. *vivax, P. malariae*, *P. ovale* as well as the immature gametocytes of *P. falciparum*. Chloroquine also has a significant anti-pyretic and anti-inflammatory effect when used to treat *P. vivax* infections, and thus it may still remain useful even when resistance is more widespread.

Amodiaquine is a 4-aminoquinoline anti-malarial drug similar in structure and mechanism of action to chloroquine. Amodiaquine has tended to be administered in areas of chloroquine resistance while some patients prefer its tendency to cause less itching than chloroquine. Amodiaquine is now available in a combined formulation with artesunate (ASAQ) and is among the artemisinin-combination therapies recommended by the World Health Organisation. Combination with sulfadoxine-pyrimethamine is no longer recommended (WHO, 2010).

2.5.2.2 Sulfadoxine-pyrimethamine

Sulfadoxine and sulfamethoxypyridazine are specific inhibitors of the enzyme dihydropteroate synthetase in the tetrahydrofolate synthesis pathway of malaria parasites. They are structural analogs of p-aminobenzoic acid (PABA) and compete with PABA to block its conversion to dihydrofolic acid. Sulfonamides act on the schizont stages of the erythrocytic (asexual) cycle. When administered alone sulfonamides are not efficacious in treating malaria but co-administration with the antifolate pyrimethamine, most commonly as fixed-dose sulfadoxine-pyrimethamine (Fansidar), produces synergistic effects sufficient to cure sensitive strains of malaria. Pyrimethamine is used in the treatment of uncomplicated malaria. It is particularly useful in cases of chloroquine-resistant *P. falciparum* strains when combined with sulfadoxine. It acts by inhibiting dihydrofolate reductase in the parasite thus preventing the biosynthesis of purines and pyrimidines, thereby halting the processes of DNA replication, cell division and

reproduction. It acts primarily on the schizonts during the erythrocytic phase, and nowadays is only used in concert with a sulfonamide.

2.5.2.3 Artemisinin and derivatives

Artemisinin is a Chinese herb (qinghaosu) that has been used in the treatment of fevers for over 1,000 years (Gomes *et al*., 2008) thus predating the use of Quinine in the western world. It is derived from the plant *Artemisia annua*, with the first documentation as a successful therapeutic agent in the treatment of malaria is in 340 AD by Ge Hong. The active compound was isolated first in 1971 and named artemisinin. It is a sesquiterpene lactone with a chemically rare peroxide bridge linkage. It is this that is thought to be responsible for the majority of its anti-malarial action, although the target within the parasite remains controversial. At present it is strictly controlled under WHO guidelines as it has proven to be effective against all forms of multi-drug resistant *P. falciparu*m, thus every care is taken to ensure compliance and adherence together with other behaviors associated with the development of resistance. It is also only given in combination with other antimalarials. The combinations of drugs currently prescribed can be divided into two categories: non-artemisinin-based combinations and artemisinin based combinations.

2.5.3 Malaria prevention and chemoprohylaxis

Malaria prevention strategies enhance the effectiveness of malaria control. Bite prevention measures include mosquito and insect repellents that can be directly applied to skin. This form of mosquito repellent is slowly replacing indoor residual spraying, which is considered to have high levels of toxicity (WHO, 2010). Further additions to preventive care are sanctions on blood transfusions. Once the malaria parasite enters the erythrocytic stage, it can adversely affect blood cells, making it possible to contract the parasite through infected blood. An experimental

approach involves preventing the parasite from binding with red blood cells by blocking calcium signaling between the parasite and the host cell. Erythrocyte-binding-like proteins (EBLs) and reticulocyte-binding protein homologues (RHs) are both used by specialized *P. falciparum* organelles known as rhoptries and micronemes to bind with the host cell. Disrupting the binding process can stop the parasite. This is known as disruptive chemoprophylaxis.

Suppressive chemoprophylactic drugs such as doxycycline and proguanil act in such a way that they are effective at killing the malaria parasite once it enters the erythrocytic stage (blood stage) of its life cycle, and therefore have no effect until the liver stage is complete. That is why these prophylactics must continue to be taken for four weeks after leaving the area of risk. Causal prophylactics target not only the blood stages of malaria, but the initial liver stage as well. This means that the user can stop taking the drug seven days after leaving the area of risk. Malarone and primaquine are the only causal prophylactics in current use (Gao *et al*., 2013).

2.6 ANTIMALARIAL DRUG RESISTANCE

The extensive deployment of these antimalarial drugs has caused immense selection pressure on human malaria parasites to evolve mechanisms of resistance. The emergence of resistance, particularly in *P. falciparum*, has been a major contributor to the global resurgence of malaria (Marsh, 1998).

2.6.1 Chloroquine resistance

Chloroquine has for decades been the primary chemotherapeutic means of malaria treatment and control (Ridley, 2002). This safe and inexpensive 4-aminoquinoline compound accumulates

inside the digestive vacuole of the infected red blood cell, where it is believed to form complexes with toxic haeme moieties and interfere with detoxification mechanisms that include heme sequestration into an inert pigment called hemozoin (Pagola *et al*., 2000). Chloroquine Resistance (CQR) was first reported in South-East Asia and South America and has now spread to the vast majority of malaria-endemic countries (Ridley, 2002). *Plasmodium falciparum* chloroquine resistance transporter gene (PFCRT) was identified as a candidate gene for CQR after the analysis of a genetic cross between a chloroquine-resistant clone (Dd2, Indochina) and a chloroquine-sensitive clone (HB3, Honduras) (Su *et al*., 1997). The PFCRT protein localizes to the digestive vacuole membrane and contains 10 putative transmembrane domains (Cooper *et al*., 2002). Point mutations in PFCRT, including the lysine 76 to threonine (K76T) mutation in the first predicted transmembrane domain, show an association with CQR in field isolates and clinical studies (Djimde *et al*., 2001). In response to high rates of treatment failure with chloroquine, Nigeria's Ministry of Health in 2005 replaced chloroquine with artemisinin-based combinations for the treatment of *falciparum* malaria (FMoH, 2005).

A previous report on chloroquine efficacy tracking has shown a phenomenal reduction in the prevalence of the major marker associated with chloroquine-resistant *falciparum* malaria years after the withdrawal of the drug signaling resurged efficacy (Kublin *et al*., 2003). The return of chloroquine-susceptible malaria following the removal of chloroquine drug pressure represents a re-expansion of a heterogeneous population of susceptible parasites that persisted during the period when chloroquine was used (Laufer *et al*., 2010). This has important implications on the effect of the removal of drug pressure on the evolution and ecology of drug resistance. Drug-susceptible organisms may regain predominance as long as there is a population that survives

despite prolonged drug pressure in the region that favours resistant parasites. This observation does not imply that parasites survive exposure to the drug; rather, it suggests that some parasites are not exposed to lethal drug concentrations, which supports the existence of a reservoir that is removed from drug pressure.

2.6.2 Sulphadoxine-pyrimethamine (SP) resistance

Due to safety concerns for ACT use during pregnancy, especially in the first trimester, SP has continued to be used in intermittent preventive treatment of malaria in pregnancy (IPTp) and infants (IPTi). Two or more doses of SP are administered for IPTp after the first trimester at intervals of at least one month apart. The importance of SP-IPTp in prevention of malaria in pregnancy and the resulting outcomes, such as low birth weight, abortion, premature birth, perinatal death, and maternal mortality, have been documented globally and WHO has continued to recommend SP-IPTp use (Menendez *et al*., 2010). SP resistance has however continued to rise and studies have reported reduced protection of SP-IPT programmes in areas where SP resistance is high (Griffin *et al*., 2010). SP resistance is caused by mutation on two genes, the dihydrofolate reductase (PFDHFR) and the dihydropteroate synthetase (PFDHPS) genes. Three PFDHFR mutations: N51I, C59R and S108N, referred to as the triple mutation, and the PFDHPS mutations: A437G and G540E, referred to as the double mutation, collectively form the quintuple mutations (Triglia *et al*., 1998). An additional mutation on PFDHPS 581 has been associated with high level of SP resistance and a strong predictor of SP-IPTp failure (Gesase *et al*., 2009) and in addition to the quintuple forms the sextuple mutation.

2.6.3 Resistance to artemisinin derivatives

Artemisinin resistance was reported first in western Cambodia (Noedl *et al*., 2008; Dondorp *et al*., 2009) where failure rates for artemisinin-based combination therapies are rapidly increasing and where resistance to previous first-line antimalarial drugs also first emerged. Artemisinin resistance has since spread, emerged independently, or both in other areas of mainland South-East Asia (Phyo *et al.,* 2012; Amarantunga *et al*., 2012; Kyaw *et al*., 2013). Artemisinin resistance is characterized by slow parasite clearance (Noedl *et al*., 2008) which reflects the reduced susceptibility of ring-stage parasites (Saralamba *et al*., 2011). It has recently been linked with point mutations in the propeller region of a *P. falciparum* kelch protein (Ariey *et al*., 2014).

2.7 VACCINE DEVELOPMENT

Vaccination is the most effective method of preventing infectious diseases and represents the greatest contribution of immunology to human health. The remarkable success of vaccines against polio, measles, diphtheria, tetanus, rabies and others, and the complete eradication of smallpox in humans prove the potential of this approach in reducing the global burden of infectious diseases (Sallusto *et al*., 2010). However, despite this success for some diseases, major logistical and technical challenges remain to be solved to develop efficient malaria vaccines that might potentially provide an important tool for use in malaria elimination and eradication programmes.

Research is leading to a new era of vaccine development in general and for malaria in particular. Adjuvants and vaccine delivery systems are becoming increasingly more important for the development of a new generation of vaccines, combining different types of adjuvants into

antigen-specific formulations with greater efficacy and improved vaccine formulations (Schijns and Lavelle, 2011). These new approaches offer a wide spectrum of opportunities in malaria vaccine research with direct applications for the near future. The main advantage of vaccine delivery systems is that they allow co-administration of immune-stimulants and more than one antigen into the same system; however, difficulties remain when applying them to human vaccinology.

Three types of vaccine candidate targeting different stages in the life cycle of the malaria parasite have been intensively investigated: (i) transmission-blocking vaccines (TBVs); (ii) pre-erythrocytic vaccines; and (iii) blood-stage vaccines.

2.7.1 Transmission-blocking vaccines

Transmission-blocking vaccines (TBVs) target antigens on gametes, zygotes and ookinetes to prevent parasite development in the mosquito midgut (Figure 2.3). The aim of these vaccines is to induce antibodies against the sexual-stage antigens to block ookinete-to-oocyst transition to stop the subsequent generation of infectious sporozoites (Carter *et al*., 2000). TBVs do not protect the recipient from contracting malaria, but could be helpful in preventing the spread of the disease. These vaccines are intended to protect entire/selected communities from infection. The leading vaccine candidates in this group include the *P. falciparum* ookinete surface antigens Pfs25 (Arakawa *et al*., 2005) and Pfs28 and their *P. vivax* homologues Pvs25 (Table 2.1) and Pvs28 (Hisaeda *et al*., 2000). To improve the immunogenicity, Pfs25 was expressed as a recombinant protein that was chemically cross-linked to Exo-Protein A and delivered as a nanoparticle. This enhanced the immunogenicity of the vaccine in mice, and it is currently undergoing Phase I trials

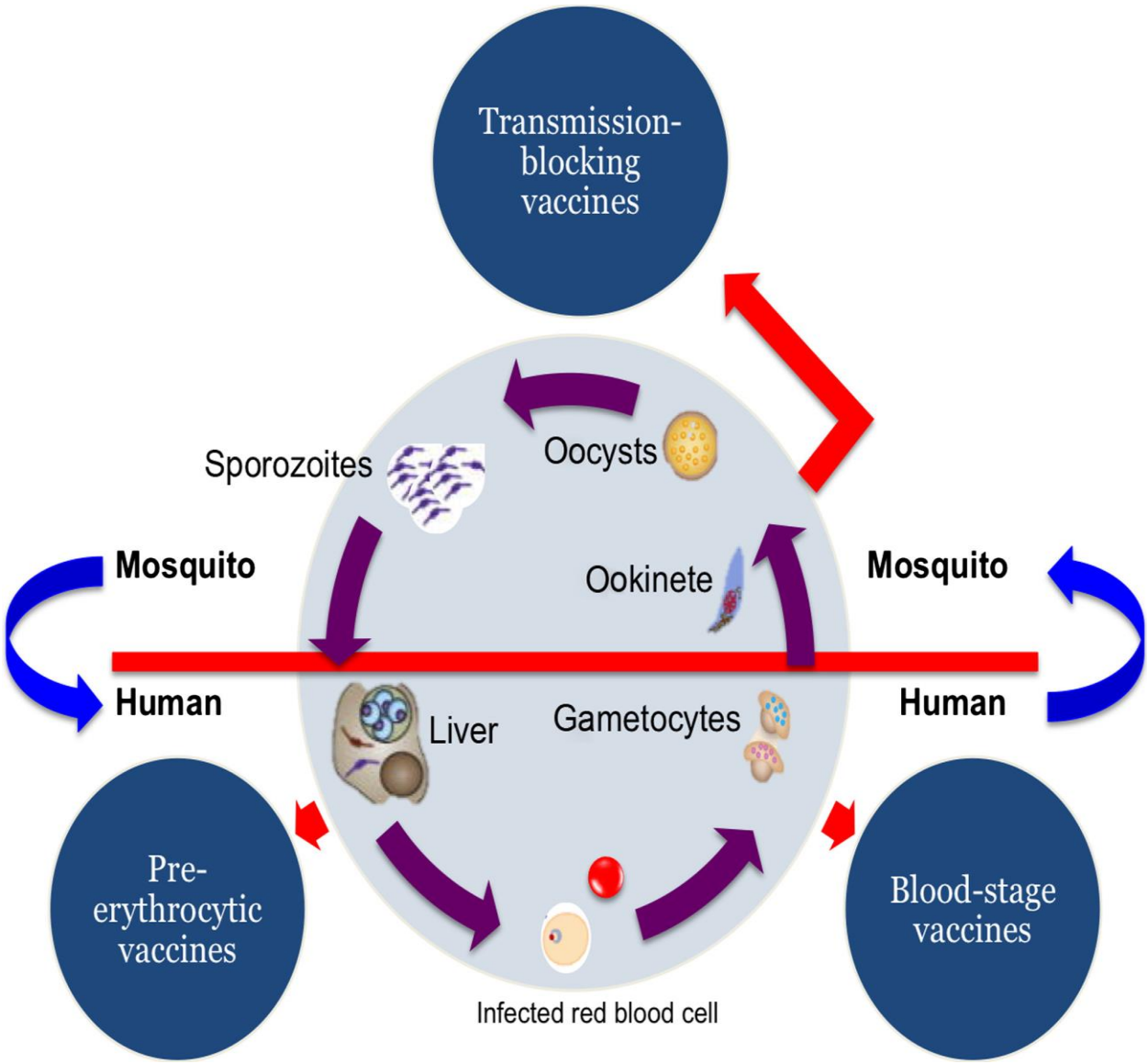

Figure 2.3: Target sites in the malaria life cycle that could be interrupted by vaccines (Carter *et al*., 2000)

Table 2.1: The major malaria vaccines that have undergone clinical trials (WHO, 2015)

Vaccine	Phase 1a	Phase 1b	Phase 2a	Phase 2b	Phase 3	Trial sponsor
Pre-erythrocytic						
RTS,S/AS 01E	+	+	+	+	+	Glaxosmith Kline, Belgium
PfCelTOS FMPO 12	+	-	-	-	-	US Army Medical Research and Materiel Command
CSVAC	+	-	-	-	-	University of Oxford
PfSpz	+	+	+	-	-	Sanaria Inc, USA
Blood-stage						
AMA-1	+	-	-	-	-	University of Oxford
BSAM-2 Alhydrogel + CPG 7909	+	-	-	-	-	National Institute of Allergy and Infectious Diseases (NIAID)
EBA	+	+	-	-	-	NIAID
SE 36	+	+	-	-	-	Research Foundation for Microbial Diseases of Osaka University, Japan
MSP3	+	+	+	+	-	African Malaria Network (AMANET)
Gmz 2	+	+	+	+	-	AMANET
Transmission-blocking						
Pfs25-EPA	+	-	-	-	-	NIAID

Key: +: Passed the stage; - : Did not pass the stage

in humans (Arevalo-Herrera *et al*., 2005). Because the antigens are never naturally presented to the human immune system, one of the potential limitations of the TBV approach is that the absence of natural boosting following immunization might limit efficacy (Okie, 2005). Nevertheless, TBVs could be important tools for a malaria elimination and eradication programme, for prevention of transmission of the disease.

2.7.2 Pre-erythrocytic stage vaccine

Pre-erythrocytic or liver-stage vaccines are designed to prevent malaria in the human host. The liver stage of *P. falciparum* is an attractive therapeutic target for the development of both antimalarial drugs and vaccines, as it provides an opportunity to interrupt the life cycle of the parasite at a critical early stage (March *et al*., 2013). However, an efficient liver-stage vaccine must be 100% effective to protect humans with no natural immunity. Such vaccines include those containing whole killed sporozoites and those based on antigenic portions of the circumsporozoite proteins (Okie, 2005). A landmark finding that set the standards for immunological protection against malaria infection was established by immunization with irradiated sporozoites (Nussenzweig *et al*., 1967).

2.7.3 Blood-stage vaccines

Blood-stage vaccines are designed to elicit anti-invasion and anti-disease responses (Moorthy *et al*., 2004). The underlying principle of these strategies is that if a vaccine could block the invasion of erythrocytes by merozoites, it would prevent malarial disease. Several blood-stage antigens have gone through clinical trials: MSP2 (Genton *et al*., 2002) and MSP3 (Audran *et al*., 2005), Apical Membrane Antigen 1 (AMA1) (Sagara *et al*., 2009), Glutamate-Rich Protein (GLURP) (Esen *et al*., 2009), merozoite surface protein (MSP) 1 (Ogutu *et al*., 2009),

Erythrocyte-Binding Antigen-175 (EBA-175) (El Sahly *et al*., 2010) and Serine Repeat Antigen 5 (SERA5) (Palacpac *et al*., 2013). All these antigens are highly expressed on the surface of merozoites. Of note, AMA1 and MSP-1 did not demonstrate efficacy in African children (Sagara *et al*., 2009; Ogutu *et al*., 2009), probably due to the highly polymorphic nature of the vaccine structures (Takala *et al.*, 2009).

Efforts to enhance the vaccine efficacy of AMA1 and MSP-1 with novel adjuvants, using viral vector prime-boost strategies or by combining AMA1 and MSP-1 have been investigated (Sagara *et al*., 2009; Ellis *et al*., 2010). The extensive genetic diversity of the parasite and the selective pressure exerted by the host's immune response are major factors to be considered in the development of effective blood-stage vaccines (Takala and Plowe, 2009). Addressing genetic polymorphism constitutes an important issue to be explored with regard to this group of vaccines. It has been suggested that efforts would be emphasized on antigens or constructs inducing cross-reactive immune responses, which could cover genetic diversity (Takala and Plowe, 2009).

Increasing research has opended the space for new antigens with great potential as blood-stage vaccine candidates to be discovered, for example, *P. falciparum* reticulocyte-binding protein homologue 5 (PfRH5) has been shown to induce inhibitory antibodies that are effective across common PfRH5 genetic variants (Bustamante *et al*., 2013). In addition, Rhoptry-Associated Leucine zipper-like Protein 1 (RALP1), which plays an important role during merozoite invasion into erythrocytes, was recognized by malaria-immune serum samples from Mali and Thailand,

suggesting the potential of this protein as a blood-stage vaccine candidate (Bustamente *et al*., 2013).

2.7.4 THE RTS, S VACCINE

The most effective malaria vaccine tested is RTS, S, a hybrid protein particle, formulated in a multi-component adjuvant named AS01. RTS, S results from collaboration, commenced in the 1980s, between the Walter Reed Army Institute of Research in the USA and GSK Biologicals, then SmithKline Beecham (Ballou and Cahill, 2007). The level of efficacy achieved by RTS, S in challenge studies was a clear breakthrough for the field and has yet to be exceeded by any sub-unit vaccine candidate. RTS, S has progressed through a series of phase I, II and III clinical trials in several African countries, involving age de-escalation from adults to infants and various efficacy assessments. These provide clear evidence that in many different epidemiological settings, RTS, S can reduce the rate of acquisition of clinical malaria by 30–50% (Tinto *et al*., 2015). The endpoint most widely accepted as a semi-standardized efficacy measure is the reduction in clinical cases (or first episodes) of malaria during the first 12 months of follow-up, a measure alluded to in the Malaria Vaccine Technology Roadmap. By this measure, the result with AS01 adjuvant formulation demonstrated an efficacy of 39 per cent in East African children (Tinto *et al*., 2015).

The breakthroughs notwithstanding, identification of malaria vaccine candidates together with the understanding of the pathogen disease mechanism and host immune response interactions has been a major challenge. This is because of the antigenic variability of the parasite. Moreover, *P. falciparum* presents numerous antigens that could feasibly be targets of protective responses.

However, such antigens are most often polymorphic, and even exhibit clonal variation through differential multigene expression (Amambua-Ngwa *et al*., 2012).

2.8 PATTERNS AND MECHANISMS OF GENETIC DIVERSITY

Many recent vaccine development projects have faced the problem of antigenic variability. Some recently developed vaccines-notably, vaccines against the encapsulated nasopharyngeal bacteria, *Haemophilus influenzae, Streptococcus pneumoniae, Neisseria meningitides* – have approached the problem of antigenic variability by multivalent formulations that raise immunity individually targeted against one or more of the most clinically significant antigenic (capsular) types. Such approaches have been largely successful, but in the case of *S. pneumoniae*, they have led to the selection of vaccine escape variants, namely strains carrying capsular types that are not included in the vaccine (Kyaw *et al*., 2013). Indeed, for many of these pathogens, antigenic variability is so pervasive that no single protective antigen has been identified (as it was for the encapsulated bacteria), widening the scope of the challenge. It is tempting to speculate that vaccine development is most difficult for pathogens that provoke the least effective immune responses following natural infection, and that these pathogens, in turn, are disproportionately prone to significant antigenic variability.

This suggests the importance of understanding the mechanisms underlying antigenic variation, in order to both design vaccines that may overcome the problem, and predict how pathogen populations will respond to the selective pressure imposed by vaccines that target particular antigenic types. Antigenic variability is also of considerable interest from the perspective of basic population biology. Understanding the forces that maintain diversity in natural populations

is a central question of both ecology and population genetics, and pathogen populations form excellent model systems for investigating these principles both theoretically and empirically. Antigenic variability on microbial populations takes many forms as pointed out by Gomes *et al*. (2008). They include:

Pattern1: Little variation within a single host, but extensive population-wide variation that is consistent in space and time.

Pattern 2: Little variation within a single host, but extensive population-wide variation that can change in space and time (*N. meningitidis*).

Pattern 3: Little variation within a single host, and little variation in the global population at a single time, but rapid variation over time on a scale of years (influenza A virus).

Pattern 4: Extensive replacement of dominant types over time within a single host, with extensive and growing standing diversity in the global population (human immunodeficiency virus, type 1).

Pattern 5: since little or no antigenic variability is known, naturally acquired immunity and vaccine-induced immunity are universal or nearly so.

If strains can be considered to exist in a strain space, in which distance corresponds to antigenic dissimilarity, the process of appearance and spread will tend to lead to clustering in strain space, because novel types similar to the most prevalent current types will constantly be generated, but types nearby an existing prevalent type will be at a disadvantage due to strong cross-immunity with the existing type. Therefore, highly prevalent types can exist only at some distance from one another, in strain space and/or time (Gomes *et al*., 2008). This effect can, under certain conditions, maintain strain structure-collections of strains with non-overlapping sets of alleles at

antigenic sites-despite recombination, because recombinants falling between non-overlapping sets will be at a disadvantage (Gupta *et al*., 1999). The group of strains present in a population may be stable or changing over time. Rapid strain turnover is promoted by higher transmission (Gomes *et al*., 2008), weaker cross-immunity (Gomes *et al*., 2008) and short-lived infections (Gupta *et al*., 1999). In the case of multilocus genotypes that determine strain structure, Gupta *et al*. (1999) found a more complex relationship with cross-immunity, in which turnover (either cyclic or chaotic) is greatest at intermediate levels of cross-immunity, with lower levels giving unstructured stable coexistence of many types and high levels of cross-immunity promoting coexistence of non-overlapping strain types.

Total pathogen prevalence is increased for pathogens having higher transmissibility and those with longer duration or less cross-immunity (Gupta and Galvani, 1999; Abu-Raddad and Ferguson, 2004). The standing diversity at any given time is greater for more transmissible pathogens and weaker cross-immunity and longer infectious durations (Abu-Raddad and Ferguson, 2004).

2.9 ANTIGENIC VARIATION AND GENETIC DIVERSITY OF PLASMODIUM FALCIPARUM

One of the most important factors for the survival of an organism is the ability to adapt to the surrounding environment. Infectious agents have developed a range of mechanisms allowing them to interact with their hosts through surface-expressed molecules while protecting themselves from the immune response. One specific mechanism is antigenic variation, when antigens undergo variations randomly or programmed, the first of which is through DNA

alterations introduced by, (1) errors in the DNA or RNA replication and repair, (2) recombination between genes, or (3) re-assortment of gene segments. Programmed variation, also called the rue antigenic variation, is characterised by a family of genes (paralogous genes) encoding proteins of similar structure and function, and the ability to express only one of these genes at the time. In this way the organism can alter the protein variant expressed from time to time (Borst, 2003).

It has been shown that diverse malaria isolates differed in their virulence patterns, by using malaria infections as treatment in syphilis patients (James *et al*., 1932). Further experiments demonstrated that immunity to malaria was strain-specific, although it was not until 1983 that clonal antigenic variation was proven in *Plasmodium falciparum* (Hommel *et al*., 1983). One appealing idea, although less likely, is that the switching is induced by some undefined signalling when anti-variant antibodies recognise and bind to a Variant Surface Antigen (VSA) expressed on the parasite infected cell (Brown, 1973). This would be an economic way for the parasite to use its VSA repertoire, since no new variant has to be produced until there are antibodies recognizing the already present one. However, there is no evidence so far supporting such a signalling mechanism. A more probable hypothesis is that there is a frequent ongoing switching in a parasite population, expressing a dominant VSA repertoire on majority of the infected cells and one or few others on a small subpopulation of cells (Figure 2.4). The outgrowth of one of these subpopulations would occur when opsonising antibodies to the parasites expressing the dominant VSAs are eliminated by clearance in the spleen. This is supported by studies in vitro showing that switching of parasite populations occurs, at rates as high as 2-4%, even in the absence of recognising antibodies (Kaviratne *et al.*, 2003).

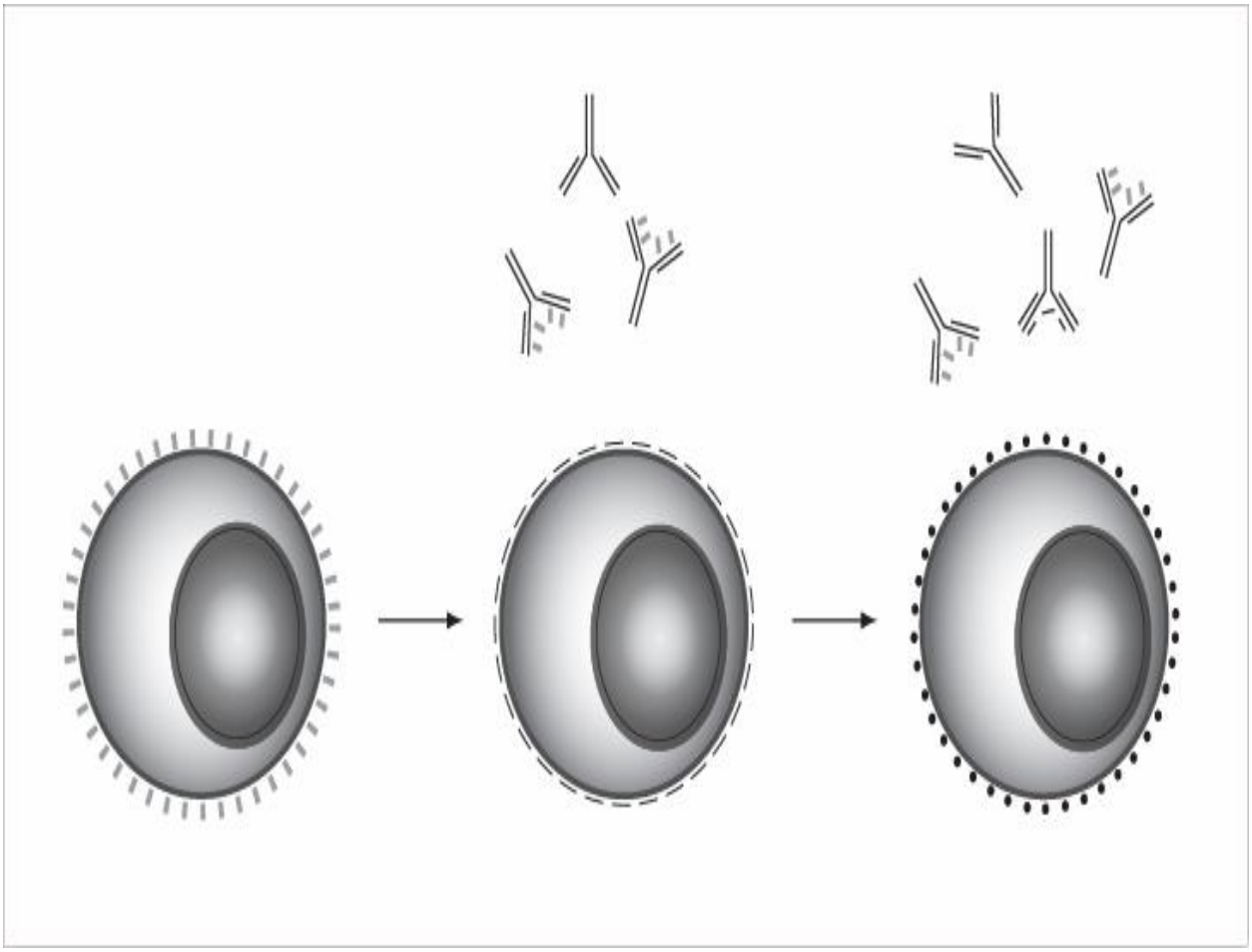

Figure 2.4: Antigenic variation in *P. falciparum.* (The repertoire of variable surface antigens (VSA) expressed on the infected erythrocyte surface can switch from one generation to another.

When analysing serum from malaria patients living in endemic regions, there is a strong correlation between the immunity of the individuals and the recognition of a broader range of different VSAs (Kaviratne *et al.,* 2003). This immunity is also associated with protection from severe disease symptoms, even though the patient might carry parasites in their circulation. The first evidence of protection from disease by the surface recognising antibodies was done in rosette disrupting assays, using serum from children with mild and severe malaria (Treutiger *et al*., 1992). A number of studies have suggested that this is the case, as association has been found both with agglutination assays and immunofluorescence, although the picture has become slightly more complex. Protective IgG in malaria have the VSAs as their main target, and the association between antibody titres to these and the level of protection is clear (Hviid, 2005). However, there is also accumulating data suggesting that there are a number of VSAs that are more common and highly recognised independently of the disease outcome (Nielsen *et al.*, 2002; Ofori *et al.*, 2002).

Plasmodium falciparum Erythrocyte Membrane Protein 1 (PfEMP1), a large protein >200 kDa, was first identified on the parasite surface after radio-iodination of intact cells followed by electrophoretic analysis of soluble and insoluble polypeptides. The size of this protein was found to vary between different parasite strains investigated and to be removable from the parasite surface by mild trypsin digestion (Ofori *et al*., 2002).

Furthermore, it was immunoprecipitated in a strain specific manner and thereby defined as a VSA (Leech *et al.*, 1984). PfEMP1 is encoded by the var gene family, with approximately 60 gene copies per haploid genome distributed over all 14 chromosomes (Su *et al.*, 1995). The

majority of the var genes are subtelomerically located and a few copies are distributed in the centre of the chromosomes. All consist of a highly variable exon 1 of 4-10 kb, a conserved 0.8-1.2 kb intron, and a conserved exon 2 of about 1.5 kb (Kaviratne *et al*., 2002). Exon 1 comprise a N-terminal Segment (NTS), 2-7 motifs of 300-400 amino acids with sequence similarity to the *P. vivax* Duffy Binding Ligand (DBL), 1-3 Cystein Rich Interdomain Regions (CIDR), and in most variants a C2 domain. It ends with a transmembrane spanning region of 25-30 amino acids (Smith *et al*., 2000). The C- terminal exon 2 is 450-500 amino acid long and >75% conserved. This domain is rich in acidic residues and is thereby called the Acidic Terminal Segment (ATS) (Su *et al.*, 1995). This is to avoid the generated VSA-specific antibodies. In early trypsin digestion experiments it was established that cytoadhesion was lost upon shaving off PfEMP1 from the infected erythrocyte (IE) surface (Leech *et al*., 1984).

This correlation between binding phenotypes and surface expression of PfEMP1 has been thoroughly investigated, and the binding to a number of receptors has been mapped to various domains of the protein. ATS is the intracellular part of PfEMP1 when inserted in the IE surface membrane and it has been shown to bind to both spectrin and actin as well as to the Knob Associated Histidine Rich Protein (KAHRP), it is hence believed to act as a protein anchor (Oh *et al*., 2000). The protein head structure DBL1α has been identified as the ligand for binding to Complement Receptor 1 (CR1), blood group A antigen, heparin, and heparan sulphate, all of which are receptors implicated to have a role in rosetting (i.e. binding of IE to uninfected erythrocytes) (Chen *et al*., 2000). CIDRα mediates binding to endothelium via CD36 and PECAM1, although the latter has also been shown to be a receptor for DBLδ. DBL2β binds to ICAM1 and DBL3 binds CSA (Chen *et al*., 2000). Other receptors that are known to be involved

in adhesion, via PfEMP1 or other yet unidentified IE ligands, are thrombospondin, VCAM, p-selectin, e-selectin and immunoglobulins (IgG, IgM) (Smith *et al*., 2000).

The repertoire of var-genes in clinical isolates varies extensively, although some functional conservation among placental binding PfEMP1 has been demonstrated (Vazquez-Macias *et al*., 2002). Polymorphism in the DBLα domain has been observed both between and within populations in endemic regions, and high diversity on genetic level has been established (Tami *et al*., 2003).

The mechanism behind this mutual selection of one gene to be expressed is still unknown, but the location and transcription direction in the genome seems to be important (Kraemer and Smith, 2003). Most var genes are located in the telomeres, which cluster together and promote silencing of these genes. However, upon relocation into transcriptionally competent perinuclear regions the genes can be activated.

Another work has identified Sir2-dependent alterations in chromatin structure and subnuclear localisation, which has a role in the regulation of var gene expression (Duraisingh *et al*., 2005). Sir2 is a histone de-acetylase that binds to repressed regions of chromosomes such as the telomeres. Disruption of this protein only de-represses a subset of subtelomeric var genes, indicating that there are other regulatory mechanisms for the rest of the gene family.

This is consistent with another observation of var subgroups, one expressed early in infection and one later (Jensen *et al*., 2004). In addition, a silencing element within the introns of the genes has been described to associate with the var promoter in a S-dependent manner, a finding consistent with the identification of a constitutively expressed var gene which lacks parts of the intron (Calderwood *et al*., 2003).

Although, another study demonstrates that a transcriptionally active var promoter is sufficient by itself to promote allelic exclusion of the parasites endogenous var gene (Voss *et al*., 2006). Taken together, there is still a lot to learn about the mutually exclusive expression of the var genes. The only thing that is definite is that it is exclusively regulated at the level of transcription, hence the idea of constitutively expressed proteins to be of regulatory importance can be discarded.

Members of a second highly polymorphic multigene family known as rif (repetitive interspersed family) were first identified from a *P. falciparum* genomic library, with a complex banding pattern in Southern blot analysis (Weber, 1988). They were confirmed at transcriptional level by Northern blots in late stage parasites, although no translation initiation codon or protein products were detected (Weber, 1988). Upon early release of data from the chromosome 2 and 3 sequencing projects, further analysis of these genes was performed suggesting that they coded for membrane proteins of a size of 27-45 kDa (Bowman *et al*., 1999). This was in the same molecular weight range as reported for the previously described rosettins, why RIFINs were first believed to be responsible for rosetting (Helmby *et al*., 1993).

PfEMP1 is the major ligand mediating rosette formation. However, detailed analysis of the surface of infected erythrocytes with radiolabelling, immunoprecipitation and trypsin cleavage revealed evidence for the rif gene product to be translated into highly variable molecules, i.e. RIFINs, which are inserted into the surface membrane and exposed on the surface (Fernandez *et al*., 1999). They are rather trypsin insensitive (>100µg/ml), hence might be responsible for the remaining adhesive capacity of IEs to PECAM1/CD31 and possible other receptors after complete shaving of PfEMP1 (<10µg/ml). RIFINs are immunogenic and widely recognised by sera from semi-immune individuals living in malaria endemic areas (Abdel-Latif *et al*., 2002). Furthermore, high titres of anti-RIFIN antibodies in patients correlate with rapid clearance of parasites upon treatment, and this reactivity is suggested to be the major anti-VSA response in natural malaria infection (Abdel-Latif *et al*., 2004).

Transcription of the rif genes peaks at 18 hours early trophozoite stages, although transcripts of some variants have been detected in later stages as well (Bozdech *et al*., 2003). Most of the rif genes are located near the var genes, within 50 kb of the telomeres, and they all share a two-exon structure. Exon 1 is coding for a short signal peptide of around 20-25 amino acids. A 140-270 bp intron separates exon 1 from exon 2, which consists of a semi-conserved 3´end followed by a hyper-variable stretch ending with a potential transmembrane domain before a short conserved intracellular 5´end. Some of the genes have a second hydrophobic putative transmembrane stretch right after the 3´ semi-conserved domain (Gardner *et al*., 2002). All rif genes have the export element (PEXEL) around 60 bp downstream exon 2 (Marti *et al*., 2004).

The inherent variability of *P. falciparum* provides multiple effective immne evasion and drug resistance mechanisms for the parasite. Many of the studies on the parasite's polymorphism have focused on variants exhibiting mutations that lead to amino acid substitutions (non-synonymous mutations) that are likely subjected to selection, such as immunogenic proteins and resistance phenotypes. Genetic diversity determines the intensity of malaria transmission, thus providing baseline data for any antimalarial drug efficacy trial and the possibility of implementing control strategies based on vaccines.

Merozoite surface proteins 1 and 2 (MSP-1 and MSP-2) are widely used to study the allelic diversity and frequency of *P. falciparum* which are most commonly correlated with the level of transmission in the area under study. In order to track the efficacy of existing antimalarial drugs, therapeutic efficacy trials are carried out over a follow-up period (WHO, 2008). Since the probability of a newly infected patient possessing a parasite genotype identical to the former infection is low (Snounou *et al*., 1998), molecular genotyping of pre-treatment (baseline) and recurrent infections enables the categorisation of recurrent parasites as recrudescent (i.e. true failure) or re-infection (i.e. successful treatment) either from pre-existing infection or a new infection from an infected mosquito bite. Therefore, investigation by comparing the genotypes of established antigenic markers (such as merozoite surface proteins, MSP 1 and 2) in patients at the time of parasite recurrence is expected to discriminate between recrudescent and new infections (Mita *et al*., 2011). However, the discriminating capability of these antigenic markers may be dependent on the extent of genetic diversity and on the frequency of each allele within the parasite population. If there is low genetic diversity of parasites in patients, an appearing

infection during the follow-up period can be incorrectly classified as recrudescent after the parasites have been cleared.

2.9.1 Merozoite Surface Proteins (MSPs)

Merozoite surface proteins are both integral and peripheral membrane protein molecules found on the surface of a merozoite. A merozoite is an early life cycle stage of a protozoan. The blood-stage cycle of the parasite is responsible for malaria pathogenesis. Intervention at this stage of the parasite's development through vaccination is likely to reduce malaria-related clinical symptoms. As a major interface between host and pathogen, the merozoite surface is an obvious target for the development of a malaria vaccine. A number of potential vaccine candidate antigens identified so far are located on or associated with the surface of the merozoite or in apical organelles (Figure 2.5).

These include merozoite surface protein 1 (MSP-1), MSP-2, MSP-3, MSP-4, MSP-5, MSP-8, RAP1/2, Apical Merozoite Antigen (AMA-1), and Erythrocyte Binding Antigen (EBA-175), which are implicated in the process of merozoite invasion of the erythrocyte (Mahanty *et al*., 2003). MSP-1 is one of the most extensively studied proteins of *P. falciparum* (Holder and Riley, 1996). It is synthesized as a ~200-kDa precursor and then processed in two steps: the primary processing step produces a complex of four fragments that are present on the merozoite surface, and the secondary processing step at invasion results in the shedding of the complex from the surface, except for the C-terminal 19-kDa domain (MSP-1_{19}), which remains anchored to the parasite surface by a glycosylphosphatidylinositol (GPI) moiety (Blackman *et al*., 1990).

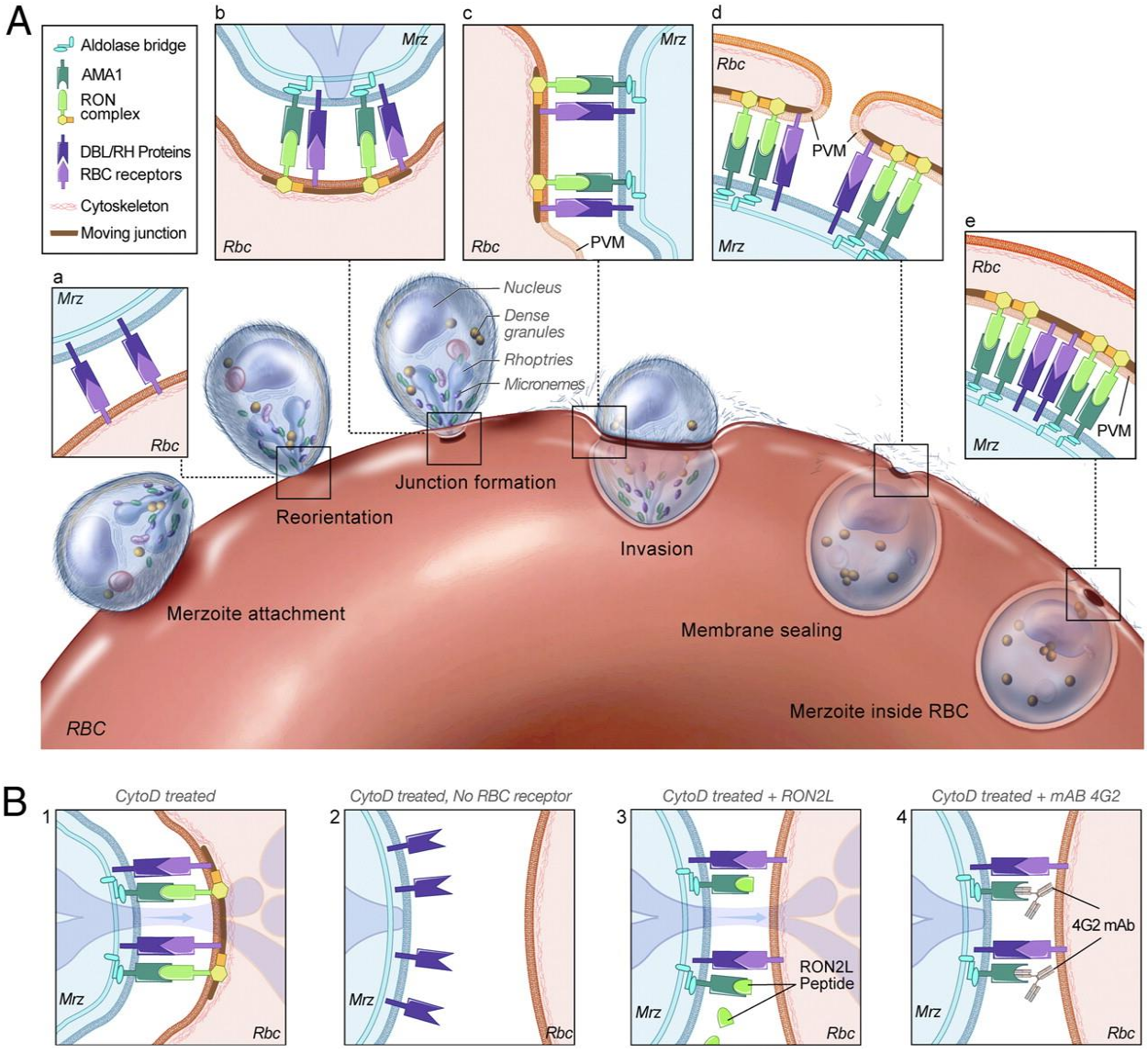

Figure 2.5: Schematic model of the steps involved in *P. falciparum* merozoite invasion. (A) Invasion of the host cell by merozoites (B) Cytochalasin-treated merozoites are able to reorient and form the junction as well as induce membrane flow (1). In the absence of the RBC receptor (e.g., GlycophorinA) the merozoite does not come into close apposition with the RBC membrane and cannot form a junction (2). In the presence of RON2L peptide (3) or AMA1-specific mAb 4G2 that competes with parasite secreted RON2 for binding AMA1 (4), junction formation is blocked. In the presence of mAb 4G2 no vacuoles are seen inside the RBC (4). However, binding of RON2L peptide to parasite AMA1 does not block rhoptry bulb secretion into the RBC (3).

(Holder and Riley, 1996)

The C-terminal 19-kDa fragment of MSP-1 is well conserved among *P. falciparum* isolates and contains two epidermal growth factor (EGF)-like domains that play a role in merozoite invasion. Substantial data from studies with *P. falciparum* MSP-1 and in vivo immunization studies of mice with *P. yoelii* and *P. chabaudi* indicate that the protective immune responses are directed against the C-terminal 19-kDa domain (Daly and Long, 1995). The inhibition of MSP-1 processing by conformation-specific antibodies (Abs) was previously proposed to be one of the possible mechanisms for the inhibition of merozoite invasion (Blackman *et al*., 1994). The MSP-1 gene has 7 variable blocks that are separated either by semi-conserved or conserved regions. Block 2, a region near the N-terminal of this gene, is the most polymorphic part of the protein and appears to be under the strongest diversifying selection within natural populations (Holder and Blackman, 1994). At present, four different allelic types of block 2 have been identified, including K1, MAD20, RO33 and MR (Takala *et al*., 2002; Happi *et al*., 2009). MSP2 is a glycosylphosphatidylinositol (GPI)-anchored protein present on the merozoite surface consisting of about 200–250 amino acids, encoded by a single exon on chromosome 2. It contains conserved N- and C-terminal (C) regions (Figure 2.6) flanking a highly polymorphic central repeat region.

A non-repeat semi-conserved dimorphic (D) region defines the two allelic families of MSP2: 3D7 and FC27 (Felger *et al*., 1997). D and C region families display low structural complexity due to the high percentage of hydrophilic residues, and are predicted and shown to represent intrinsically unstructured regions (Wright and Dyson, 1999). It has been shown that specific semi-immune Ab against MSP2 protein is predominantly cytophilic IgG3, as in other blood stage

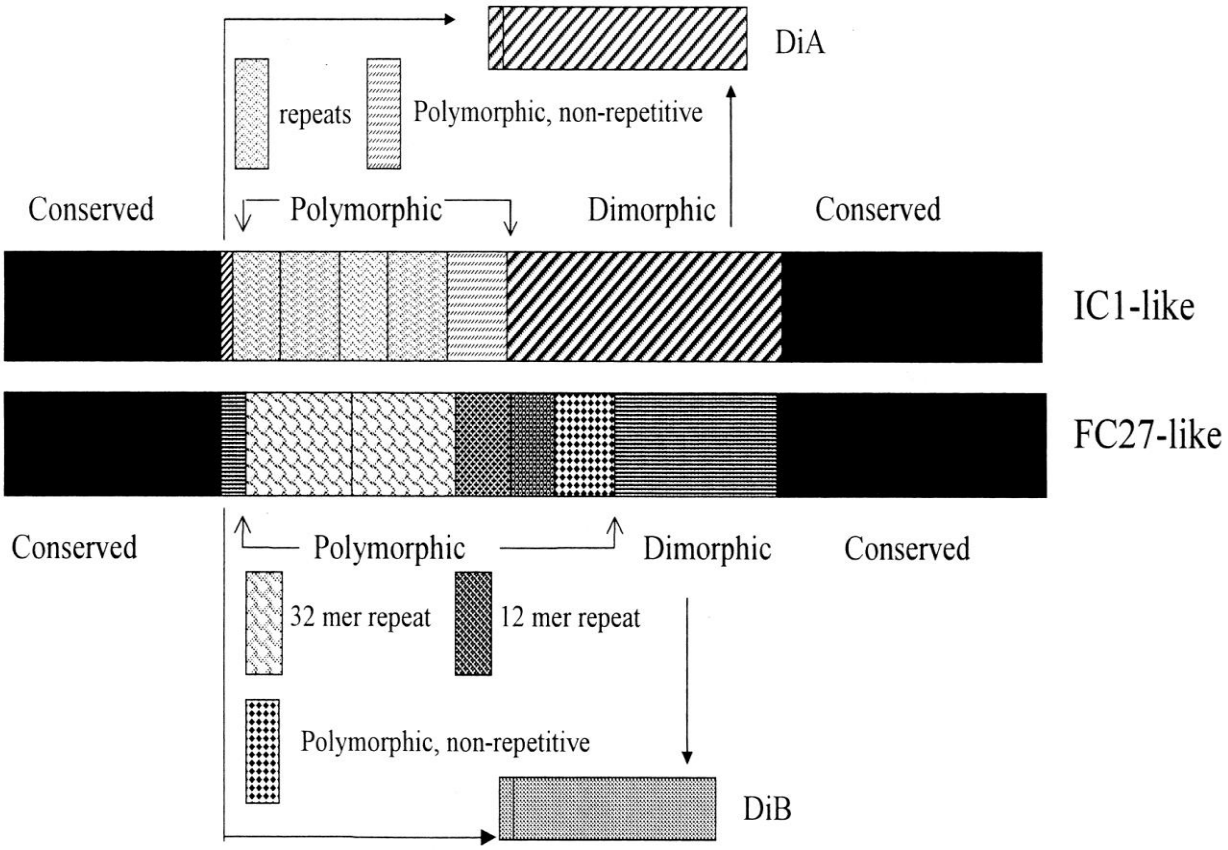

Figure 2.6: Schematic representation of merozoite surface protein-2

(Franks *et al*., 2003)

proteins (Taylor *et al*., 1995). These cytophilic (IgG1 and IgG3) Abs are thus thought to play an important role in antibody-mediated mechanisms of parasite clearance (Tebo *et al*., 2001).

Merozoite surface protein, MSP-3, has been shown to be the target of the protective immune responses in humans (Oeauvray *et al*., 2000). The PfMSP-3 protein contains three blocks of four tandem heptad repeats based on the AXXAXX motif at the N terminus, a glutamic acid-rich domain, and a putative leucine zipper sequence at the C terminus (McColl *et al*., 1994). Although a clear surface localization of PfMSP-3 is known, it lacks any transmembrane domain or glycosylphosphatidylinositol (GPI) anchor site (McColl and Anders, 1997) and is therefore considered to be loosely associated with the merozoite surface by interactions with other merozoite surface proteins. PfMSP-3 was identified as a candidate vaccine antigen by an Antibody-Dependent Cellular Inhibition (ADCI) assay using human immune sera (Oeuvray *et al*., 1994).

The potential of PfMSP-3 as a vaccine candidate was further illustrated by ADCI using mice antibodies and was further confirmed by the suppression of *P. falciparum* growth in an immune-compromised mouse after the passive transfer of human antibodies purified on MSP-3 peptides together with human monocytes (Oeuvray *et al*., 1994).

The immunization of Aotus and Saimiri monkeys with recombinant PfMSP-3 or its fragments provided protection against parasite challenge (Carvalho *et al*., 2004). A 70-amino-acid-long conserved domain of PfMSP-3, referred to here as the PfMSP-311 region, was identified as the

target of protective antibodies in human immune responses. The presence of high titers of cytophilic antibodies, IgG3, against this conserved region of MSP-3 has been correlated with protection against the parasite. In addition, immunization of humans with a synthetic peptide corresponding to this region was previously shown to induce antiparasitic antibodies that suppress parasite growth in an ADCI assay.

2.10 PLASMODIUM FALCIPARUM SURFACE MOLECULES, IMMUNITY AND POLYMORPHISMS

Polymorphisms in repeats may enable parasites to evade immune responses elicited by past exposure to diverse forms of the same antigen. In addition, tandem repeats may stimulate T-cell-independent B-cell responses (Schofield *et al*., 1987) that fail to generate memory B cells or somatic hyper-mutation, leading to antibody affinity maturation. Moreover, cross-reactive epitopes on otherwise different repetitive antigens may prevent the affinity maturation of antibodies by causing an abnormally high proportion of mutated B cells to be preserved during clonal expansion (Anders, 1986). An example of such a cross-reactive epitope is the pentapeptide VTEEI, present in repeat arrays of unrelated blood-stage (RESA and Pf332) and gametocyte antigens of *P. falciparum* (Ahlborg *et al*., 1991).

Minor amino acid diversity is created in malarial antigens by single-nucleotide replacements. These point mutations are relatively rare events (10^{-9} replacements/nucleotide/replication) that occur at any site, but they are often clustered in sequences coding for B- or T-cell epitopes of malarial antigens. At the circumsporozoite (CSP) locus of *P. falciparum*, nearly all nonsynonymous nucleotide replacements (i.e., those leading to amino acid replacements) map to

the T-cell epitopes Th2R and Th3R, located downstream of the central repeat, putatively due to the positive selection of novel polymorphisms associated with immune evasion (Conway and Polley, 2002).

In fact, amino acid replacements in the CSP T-cell epitopes Th2R and Th3R abolish their recognition by naturally exposed people (Bonelo *et al*., 2000). Amino acid replacements are found in other T-cell epitopes of malarial antigens, but their impact on T-cell recognition and immune evasion remains to be confirmed. Genetic recombination, however, accounts for most variation seen in malarial antigens, since it occurs several orders of magnitude more frequently than mutations. In addition, a single event changes several nucleotides at once. Recombination events comprise both exchanges of blocks of homologous sequences during meiosis and rearrangements in repeat arrays during both mitosis and meiosis. The extensive variation in CSP, MSP2, and many other antigens results primarily from insertions and deletions of repeat motifs (Berzins and Anders, 1999). In contrast, both exchanges of non-repetitive sequences and rearrangements within repeat arrays seem to generate new MSP-1 alleles in natural populations of *P. falciparum* (Ferreira *et al*., 2003) and *P. vivax* (Putaporntip *et al*., 2002).

Allelic dimorphism, the occurrence of only two alternate forms at a given locus, characterizes several malarial antigens. The dimorphic allelic families are not necessarily homogeneous (i.e., there is some within-family heterogeneity); members of different families, however, differ at more than 30% of the amino acid residues in variable sequences (i.e., there is extensive between-family divergence). Allelic dimorphism was originally described for the MSP-1 locus (Tanabe *et al.*, 1987) and was later found in other blood-stage antigens of *P. falciparum*, such as the

merozoite surface proteins MSP-2 and MSP-3, the serine repeat antigen (SERA), and erythrocyte-binding antigen (EBA-175), a parasite molecule that binds to sialic acid and glycophorin A during RBC invasion (Conway and Baum, 2002). The MSP-1 locus of *P. vivax* is also essentially dimorphic (Gibson *et al*., 1992).

The repetitive sequences in block 2 of the *P. falciparum* MSP-1 gene may be taken as an example: the proliferation of trinucleotide motifs and fusion with nearby triplets, followed by several nucleotide replacements, may explain the origin of the whole repertoire of 9-bp repeats of the dimorphic allelic families. Similar processes have been postulated to operate at the MSP-2 locus of *P. falciparum* (Felger *et al*., 1997). Alternatively, dimorphic alleles could arise by gene duplication followed by sequence divergence between paralogous genes as a result of gene conversion events (Hartl *et al*., 2002). This process may readily explain the origin of allelic dimorphism in EBA-175 and SERA, which are encoded by members of multigene families.

Cross-reactivity is a major factor driving the emergence and persistence of novel antigenic variants of malaria parasite antigens in human populations; new variants are theoretically selected if mutant parasites evade the host's immunity (Mckenzie *et al*., 2001). Thus, the high recombination rate typically found in repeats, coupled with the positive selection of non-cross-reacting variants, might account for most of the antigenic diversity seen in natural malaria parasite populations (Rich and Ayala, 2000).

Naturally acquired immune responses discriminate between antigenic variants belonging to different dimorphic families of *P. falciparum* MSP-1 and MSP-2 (Cavanagh *et al*., 1998) and *P. vivax* MSP-1 (Mancilla *et al*., 1994). Significantly, experimental immunization with a

multivalent vaccine prototype containing the 3D7 variant of MSP-2 partially protected humans from infection with homologous (IC1-type) but not heterologous (FC27-type) parasites (Genton *et al*., 2002). However, insertions and deletions of repeat units within dimorphic allelic families affect immune recognition and favor immune evasion: (i) narrowly specific antibodies to repetitive block 2 variants of MSP-1 are found in subjects exposed to malaria (Da Silveira *et al.*, 1999), and (ii) murine (Locher *et al*., 1996) and human (Sowa *et al*., 2001) monoclonal antibodies are able to discriminate between repeat variants within the same dimorphic allelic family. The monoclonal antibody described by Sowa *et al*. (2001), for example, discriminates between the MAD20 and HB3 variants of block 2, which differ only in the number of copies of the hexapeptide SGGSVA.

The expression of variant antigens on the surface of infected RBCs is an additional strategy of immune evasion used by all *Plasmodium* species so far studied (Paget-Mcnicol *et al*., 2012). At a first glimpse, the expression of foreign antigens on RBCs, which are devoid of major histocompatibility complex molecules and which thus represent the ideal hideout for intracellular organisms, seems a disadvantage for the parasite's survival in the presence of a competent immune system. However, this may be an intrahost mechanism used to control parasite populations, since rapid parasite multiplication would kill the host in a very short time, decreasing the chances of transmission to mosquitoes (Paget-McNicol *et al*., 2002).

2.11 SIGNATURES OF SELECTION ON PLASMODIUM FALCIPARUM

Host-pathogen interactions create selection for resistant genotypes which can either become fixed or polymorphic with different mutant alleles present within or between populations. Such interfaces between the parasite and human host’s immune system impose strong pressure and

leave selections in the genes responsible for drug resistance and/or evasion of host's immunity (Dieckman *et al*., 2002; Lipsitch *et al*., 2007; Fumagalli *et al*., 2008). A number of genes encoding drug or immune targets have been linked with natural selection in *P. falciparum* identified as signatures of directional or balancing selections (Escalante *et al*., 1998; Volkman *et al*., 2007; Mu *et al*., 2007; Weedall *et al*., 2008).

Directional selections cause adaptively important genetic variants of the parasite to increase in frequency leading to high fixation rates, the appearance of a selective sweep and reduced variability in the region flanking the selected loci. Previous analyses of microsatellites and single nucleotide polymorphisms (SNPs) have identified such selective sweeps around some known drug-resistance genes (Wootton *et al*., 2002; Hall *et al*., 2005; Jeffares *et al*., 2007; Weedall *et al*., 2008). Single nucleotide polymorphisms contribute largely to the variability. Alignment of sequences based on information obtained from multiple geographical locations has shown that synonymous mutations are scarce in most of the parasite's genes. Genes encoding proteins that are expressed on the surface of sporozoites or merozoites are, in general, more variable than housekeeping genes or genes expressed during the sexual stages and, thus, constitute a major obstacle in targeting them in the design of vaccines. Comparison of the rather selectively neutral synonymous and the more frequent non-synonymous substitutions indicate that natural selection may account for most polymorphisms observed at functional gene loci (Escalante *et al*., 1998).

The diversifying mechanisms of single non-synonymous mutations resulting in the substitution of amino acids and, as a second process, rapid intragenic recombination observed in gene segments with repeat motifs have been proposed to underlie the polymorphism in those genes

that are subject of adaptive selection (Rich and Ayala, 2000). Non-synonymous SNPs are found at high frequencies in genes that are under strong selective pressure, in particular those genes encoding proteins linked to immune evasion or drug resistance. For example, drug resistance have been found in the dihydrofolate reductase-thimidylate synthase (DHFR-TS) and the dihydropteroate synthetase (DHPS). These genes are associated with resistance to pyrimethamine and sulfadoxine, respectively (Wang *et al*., 1997). Other SNPs have been claimed of being responsible for chloroquine (CQ) resistance (Fidock *et al*., 2000).

Balancing selection brings the parasite's favoured alleles to an intermediate equilibrium, where they are maintained as genetic polymorphisms that potentially increase the variability in the genomic region flanking the selected loci. Analyses of genes encoding surface protein targets of adaptive immunity have shown signatures of balancing selection maintaining intermediate stability of different favoured alleles within populations (Conway and Polley, 2002; Volkman *et al*., 2007; Mu *et al*., 2007). Thus systematic scan for signatures of balancing selection has become imperative to prospect for new vaccine developments based on multi-allelic antigen devising. Before the genomics era, deductions on genes under balancing selection were made from candidate gene studies. These single gene analyses have yielded some remarkable success in providing information on the evolutionary history of *P. falciparum* (Anderson *et al.,* 1999; Aurrecoechea *et al*., 2009; Mu *et al*., 2010).

However, the analyses of such candidate genes have some major drawbacks. For instance, prior knowledge was needed of the genes that might have been subjected to selection which was achievable only for genes with well-defined phenotypes as against targets of selections involved

in immunity. Interpreting population evolutionary history and patterns of genetic variation at individual loci is also often confounding (Akey, 2009). Thus advances in genome sequencing methods and the growing knowledge of array-based association techniques have brought analyses of patterns of balancing selection in *P. falciparum* closer to discernment.

2.12 GENOME ANALYSIS AND PSEUDOGENES

The genome is the total genetic complement (the sum of all the genes and intergenic sequences) of a cell. The genome is largely stable but subject to changes due to very slow evolution over a time scale. The first sequenced genome of a living organism was that of *Haemophilus influenza* (Fleischmann *et al*., 1995). In eukaryotes, the major part of the genome is sequestered in the nucleus that is delimited from the surrounding cytoplasm by a porous nuclear membrane. At the gross level, the total amount of nuclear DNA varies in different organisms and is strictly species specific. A relatively small part of the genome is contained in the organelles (chloroplasts and mitochondrion), which show wide characteristic variations, differing from the type of sequences that constitute the nuclear genome. The sequences of the DNA comprising the gene are thus divided into two categories. The exons comprise of the regions that are represented in the mRNA and used to produce the protein product. The introns are missing from the mRNA. The process of gene expression involves a new step, one that does occur in bacteria. The DNA produces a copy of RNA with the exact copy of gene sequence. This RNA is only a precursor and in order to produce a functional mRNA, it undergoes splicing to remove introns (non-coding sequences).

Pseudogenes consist of sequences that are related to functional genes, but cannot be translated into functional proteins. There are some pseudogenes that have the same structure as the

functional genes that correspond to exons and introns in the same locations. It has been suggested that these genes are rendered inactive due to mutations, hence cannot be expressed.

The mutations are generally responsible for removing the signal sequences initializing transcription, for preventing splicing, or terminating transcription prematurely. This means that pseudogenes are caused by a number of deleterious mutations in the functional gene over a period of time. Pseudogenes are formed in a number of systems that include globin, immunoglobulins and histocompatibility antigens.

2.12.1 Genome mapping

Genomes of eukaryote are arranged on chromosomes. Genes that lie on the same chromosome are usually inherited together to form a linkage group. Often, maternal and paternal chromosomes exchange piece of their DNA in a process called genetic crossover. The locus of a gene on a chromosome can be evaluated using the probability of crossover events. Chromosome consists of two parts: Euchromatin and heterochromatin. Euchromatin contains most of the genes and this is the part of the genome that is actually sequenced. Heterochromatin contains mostly short sequences and its function is unknown. Whole genome analysis is now possible and it involves the development of genome maps. A genome map is a graphic representation that provides information about the location of genes. The two types of genome maps include genetic and physical maps. Genetic maps are based on meiotic recombination frequencies, and the distance between loci is measured in centiMorgans (cM). Genetic mapping gives rise to a map showing the positions of the genes relative to each other, and relative to the ends and centre of the chromosomes.

Linkage maps show the arrangement of genes and genetic markers along the chromosome as calculated by the frequency with which they are inherited together. The linkage map can be constructed by observing how frequently two markers are inherited together. This is based on the presumption that two markers located closely on the same chromosome will tend to be passed together from parent to offspring. During the normal process, DNA strands occasionally break and rejoin in different places on the same chromosome or on the homologous chromosome in a process known as meiotic recombination. Meiotic recombination can result in the separation of two markers originally on the same chromosome. The linked genes tend to remain together and are less likely to participate in a recombination event. Meiotic recombination frequency thus provides an estimate of the distance between two markers. Genetic markers must be polymorphic for them to be useful in linkage analysis. A variety of DNA polymorphic DNA markers are now available that can be used for linkage studies. These include Restriction Fragment Length Polymorphisms (RFLPs), Variable Number Tandem Repeats (VNTRs), Sequence Tagged Sites (STS) and Expressed Sequence Tags (ESTs).

2.12.2 Restriction Fragment Length Polymorphisms (RFLP)

Restriction fragment length polymorphism is a technique in which organisms are differentiated by analysis of patterns derived from cleavage of their DNA. If two organisms are differentiated by analysis of patterns derived from cleavage of their DNA. If two organisms differ in the distance between sites of cleavage of a particular restriction endonuclease, the length of the fragments produced will differ when the DNA is digested with a restriction enzyme. RFLP has been used to analyze infection dynamics of individual clones of *P. falciparum* (Smith *et al*., 2000).

2.12.3 Variable Number Tandem Repeats (VNTRs)

A variable number tandem repeat is a location in a genome where a short nucleotide sequence is organized as a tandem repeat. These can be found on many chromosomes, and often show variations in length between individuals. Each variant acts as an inherited allele, allowing them to be used for personal or parental identification. Variable number tandem repeat polymorphisms are appreciated as a type of copy number variation (Conrad *et al*., 2010) and have been surveyed in the Open Reading Frames (ORFs) of *Neisseria* spp. (Jordan *et al*., 2003), humans (O'Dushlaine *et al*., 2005), *Saccharomyces cerevisiae* (Bowen *et al*., 2005; Verstrepen *et al.*, 2005), *Aspergillus fumigatus* (Levdansky *et al.*, 2007) and *Legionella pneumophila* (Coil *et al*., 2008). The usefulness in genetics and biology research, forensics, and DNA fingerprinting has also been tested in *P. falciparum* (Tan *et al.*, 2010). Investigations of repeat structures in the *P. falciparum* genome include repeats in genes that encode antigen or surface proteins (Cowman *et al*., 1984), subtelomeric repeats including the rep20 elements (Oquendo *et al*., 1986) and microsatellite sequences (Su *et al*., 1999).

2.12.4 Sequence Tagged Sites (STS)

A sequence-tagged site is a short (200 to 500 base pair) DNA sequence that has a single occurrence in the genome and whose location and base sequence are known. STSs can be easily detected by the polymerase chain reaction (PCR) using specific primers. For this reason they are useful for constructing genetic and physical maps from sequence data reported from many different laboratories. They serve as landmarks on the developing physical map of a genome.

2.12.5 Expressed Sequence Tags (ESTs)

An expressed sequence tag is a short sub-sequence of a complementary DNA sequence. They may be used to identify gene transcripts, and are instrumental in gene discovery and gene sequence determination (Adams *et al*., 1992). Analysis of expressed sequence tags from *P. falciparum* has been documented (Chakrabarti *et al*., 1994).

2.12.6 Physical maps

While the genetic map uses crossover frequency as a basis for map distance in centiMorgans (cM), the physical map is based on the actual location of genes in the chromosome. This comprises ordered DNA clones that cover a complete chromosome and describe the chemical characteristics of the DNA molecule itself. When the DNA clones overlap each other in a contiguous array, they are called a contig. A contig map can be used to generate the complete DNA sequence of a chromosome. On a physical map, the distance between genes is measured in base pairs (bp) of DNA. The different types of physical maps include chromosomal map, expression map, radiation hybrid map, STS maps and clone maps. Chromosomal or cytogenetic map is based on the distinctive banding patterns observed by light microscopy of stained chromosomes. As chromosomal maps are based on estimates of physical distance of the markers, they are considered to be physical maps. In a chromosomal map, genes or other identifiable DNA fragments are assigned to their respective chromosomes, with distances measured in base pairs. These markers can be physically associated with particular bands (identified by staining) primarily by in situ hybridization (ISH). ISH that involves tagging the DNA marker with an observable label (e.g one that fluoresces or is radioactive). The location of the labeled probe can be detected after it binds to its complementary DNA strand in an intact chromosome.

Chromosomal mapping can be used to locate genetic markers defined by traits observable only in whole organisms. The number of base pairs within a band can only be estimated. Expression map shows the positions of expressed DNA regions (exons) on the chromosomal map. Expressed DNA regions are those transcribed into mRNA, cDNA is synthesized in the laboratory using the mRNA molecule as a template; base-pairing rules are followed. This cDNA can then be mapped to genomic regions. As they represent expressed genomic regions, cDNAs are thought to identify the parts of the genome with the most biological and medical significance. A cDNA map can provide the chromosomal location for genes whose functions are currently unknown. Sequence Tag Site (STS) is a sequence of about 200 nucleotides in length and it is hypothesized that STS occurs exactly once on the entire genome. Because of this uniqueness, STSs are often used as markers on physical maps. They are considered landmarks for locating other interesting sites. STSs by themselves may not be genetically meaningful themselves.

2.13 *PLASMODIUM FALCIPARUM* GENOME STRUCTURE

The *P. falciparum* genome is composed of 22.8 megabases (Mb) distributed among 14 chromosomes ranging in size from approximately 0.643 to 3.29 Mb. The overall (A + T) composition is 80.6%, and rises to ~90% in introns and intergenic regions. The structures of protein-encoding genes were identified. Introns are contained in 54% of *P. falciparum* genes. Fifty-two percent of the predicted gene products were detected in cell lysates prepared from several stages of the parasites life cycle by high-resolution liquid chromatography and tandem mass spectrometry (Florens *et al*., 2002), including many predicted proteins with no similarity to proteins in other organisms.

In addition, 49% of the genes overlapped (97%) identity over at least 100 nucleotides with expressed sequence tags (ESTs) derived from several life-cycle stages. As proteomics and EST studies may not represent a complete sampling of all genes expressed during the complex life cycle of the parasite, this suggests that the annotation process identified substantial portions of most genes. However, in the absence of supporting EST or protein evidence, correct prediction of the 5′ ends of genes and genes with multiple small exons is challenging, and the gene models should be regarded as preliminary. Additional ESTs and full length complementary DNA sequences (Watanabe *et al*., 2001), are required for the development of better training sets for gene-finding programs and the verification of the predicted genes. The nuclear genome contains a full set of transfer RNA (tRNA) ligase genes and 43 tRNAs were identified to bind all codons except TGT and TGC, coding for Cys; it is possible that these tRNAs are located within the currently unsequenced regions. All codons ending in C and T appear to be read by single tRNAs with a G in the first position, which is likely to read both codons via G:U wobble. Each anticodon occurs only once except for methionine (CAT), for which there are two copies, one for translation initiation and one for internal methionines, and the glycine (CCT) anticodon, which occurs twice. A putative selenocysteine lyase is present, which may provide selenium for synthesis of selenoproteins. Increased growth has been observed in selenium-supplemented *Plasmodium* culture (Gamain *et al.,* 1996). The genome does not contain long tandem repeated arrays of ribosomal RNA (rRNA) genes. Instead, it contains several single 18S-5.8S-28S rRNA units distributed on different chromosomes.

The sequence encoded by an rRNA gene in one unit differs from the sequence of the corresponding rRNA in other units. Furthermore, the expression of each rRNA unit is

developmentally regulated, resulting in the expression of a different set of rRNAs at different stages of the parasite life cycle (Li *et al*., 1994). It is likely that by changing the properties of its ribosomes the parasite is able to alter the rate of translation, either globally or of specific messenger RNAs (mRNAs), thereby changing the rate of cell growth or altering patterns of cell development. The two types of rRNA genes previously described in *P. falciparum* are the S-type, expressed primarily in the mosquito vector, and the A-type, expressed primarily in the human host. Two copies of the S-type rRNA genes are located on chromosomes 11 and 13, and two copies of the A-type genes are located on chromosomes 5 and 7. In addition, chromosome 1 contains a third, previously uncharacterized, rRNA unit that encodes 18S and 5.8S rRNAs that are almost identical to the S-type genes on chromosomes 11 and 13, but has a significantly divergent28S rRNA gene (65% identity to the A-type and 75% identity to the S-type). The expression profiles of these genes are unknown. Chromosome 8 also contains two unusual rRNA gene units that contain 5.8S and 28S rRNA genes but do not encode 18S rRNAs; it is unknown whether these genes are functional. The sequences of the 18S and 28S rRNA genes on chromosome 7 and the 28S rRNA gene on chromosome 8 are incomplete as they reside at contig ends. The 5S rRNA is encoded by three identical tandemly arrayed genes on chromosome 14. *P. falciparum* chromosomes vary considerably in length, with most of the variation occurring in the subtelomeric regions. Field isolates, even those from individuals residing in a single population exhibit extensive size polymorphism that is thought to be due to recombination events between different parasite clones during meiosis in the mosquito (Hinterberg *et al*., 1994).

Chromosome size variation is also observed in cultures of erythrocytic parasites, but is due to chromosome breakage and healing events and not to meiotic recombination. Subtelomeric

deletions often extend well into the chromosome, and in some cases alter the cell adhesion properties of the parasite owing to the loss of the gene(s) encoding adhesion molecules. Because many genes involved in antigenic variation are located in the subtelomeric structure and functional properties is essential for elucidation of the mechanisms underlying the generation of antigenic diversity. The subtelomeric regions of the chromosomes display a striking degree of conservation within the genome that is probably due to inter-chromosomal exchange of subtelomeric regions (Gardener *et al*., 2002).

In common with other organisms, highly variable gene families are clustered towards the telomeres. *Plasmodium falciparum* contains three such families termed *P. falciparum* erythrocyte membrane protein 1 var gene (PfEMP1), repetitive interspersed family (RIFIN) and Sub-Telomeric Variable Open Reading frame (STEVOR) (Kyes *et al.*, 1999). The reference 3D7 genome contains 59 var, 149 RIFIN and 28 STEVOR genes. The var genes code for proteins which are exported to the surface of infected red blood cells where they mediate adherence to host endothelial receptors, resulting in sequestration of infected cells in a variety of organs. These and other adherence properties are important virulence factors that contribute to the development of severe disease. RIFINs are also expressed on the surface of infected red cells and undergo antigenic variation (Kyes *et al*., 1999). Proteins encoded by STEVOR genes show sequence similarity to RIFINs, but they are less polymorphic. PfEMP1 proteins are targets of the host protective antibody response (Bull *et al*., 1998), but transcriptional switching between var genes permits antigenic infection and transmission. Products of the var gene family are thus central to the pathogenesis of malaria and to the induction of protective immunity.

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 THE STUDY AREAS

Participants presenting with symptoms of malaria were recruited at three health facilities State Specialist Hospital, St Kizito's Catholic Hospital and Ajara Primary Health Centre, respectively representing three localities in southwestern Nigeria - Aramoko-Ekiti (AMK); Lekki (LEK) and Badagry (BDG) (Figure 3.1). The choice of AMK followed its molecular epidemiological importance as a rural inland community in Ekiti State with approximately 74, 491 people (NPC, 2006). It is an upland zone rising over 250 meters above sea level with characteristic Guinea Forest-Savannah Mosaic Vegetation. The geographical coordinates are 7^0 43′ 0″ North, 5^0 3′ 0″ East. LEK is an urban settlement in Lagos State with approximately 117, 793 residents (NPC, 2006). It is a coastal area which lies between 6^0 25′ 0″North and 4^0 6′ 0″East. BDG is a peri-urban transitory community in Lagos State bordering Nigeria and Republic of Benin. It has a population of 241, 093 (NPC, 2006) and lies between 6^0 25′ 0″North and 2^0 53′ 0″East. Year-round malaria transmission occurs in all study areas and peaks during the annual rainy season from March to October.

3.2 ETHICAL CLEARANCE

Informed consent was obtained from individuals and parental consent and/or assent from under-16 participants prior to recruitment. Ethical approval with reference number IRB/12/209 was obtained from the Nigerian Institute of Medical Research (NIMR) ethics committee (Appendix 1).

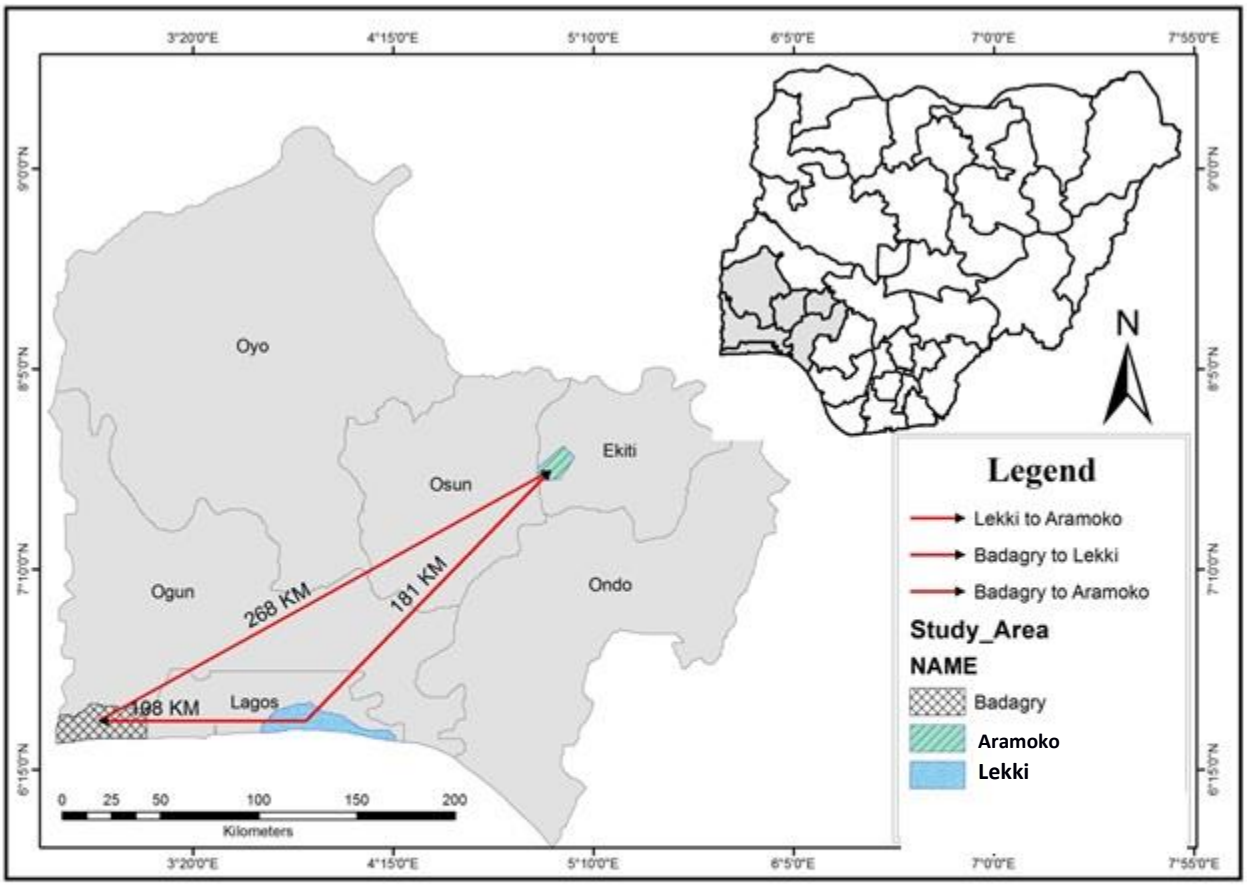

Figure 3.1: Map of southwestern Nigeria showing the study areas and the geographic distances between them

3.3 EXPERIMENTAL DESIGN AND SAMPLE COLLECTION

This is a cross-sectional study which involved sample collection from participants 2 years and above seeking care for uncomplicated malaria at selected health centres in the study areas from March to October, 2013. Samples were recruited from both urban and rural settings to inquire into population differentiation, given the variance in access to drugs and other interventions (Omole and Onademuren, 2010). Up to 5ml of venous blood was collected from each participant in sterile Ethylene Diamine Tetracetic Acid (EDTA) – coated vacutainer tubes and blood spots were made on 3mm Whatmann filter paper (Whatmann, England) and left to dry after which participants were treated with artemether-lumefantrine (20-120mg) according to national policies (FMoH, 2005). Collected samples were transported in an ice-cold container to the laboratory for further analyses.

The sample size was calculated using World Health Organization's statistical method (WHO, 2003). The assumed prevalence rate was 24%. At a confidence level of 95% and a precision around the estimate of 5%, a minimum of 196 patients were included. With a 20% increase to allow for those who refused to consent to test, 240 patients were enrolled in the study per area:

$$N = \frac{z^2 p(1-p)}{d^2} \quad \text{.....} \quad \textbf{Equation 3.1}$$

N = Average number of *P. falciparum*-infected patients attending the hospitals within the period of one year. Given the sample population of *P. falciparum*-infected patients attending a particular clinic within the period of one year (N) to be 400,

$$nf = \frac{280}{1+\frac{280}{400}}$$

$$nf \approx 165 \text{ samples}$$

3.4 RECRUITMENT CRITERIA

Blood samples were collected from participants 2 years and above seeking care for symptomatic malaria at selected health centres in the study localities from February to October, 2013. All recruited malaria cases had a temperature of > 37.5^0C on presentation or history of fever in the previous 48 hours, and a minimum of 5000 *P. falciparum* parasites/ul estimated by thick film examination.

3.5 MICROSCOPY

Thick and thin blood films were prepared on microscope slides and the thin films were properly fixed with methanol and left to dry. The prepared slides were then stained with Giemsa and examined for the presence of malaria parasites under 100x oil-immersed lens. A minimum of 2 – 200 high-power fields in each blood film were examined depending on parasitaemia levels and the negative results for parasites were declared with 200 high-power fields examined in blood films (Greenwood and Armstrong, 1991). Thin blood smear confirmed each infection as *P. falciparum* only. Parasitaemia was calculated using the formula:

$$\frac{Numberofparasites\ counted}{Numberofleucocytes} x8000\frac{parasites}{ul}ofblood \quad \text{.....Equation 3.2}$$

3.6 DNA EXTRACTION FROM BLOOD SPOTS

Three punched out discs from blood spotted on 3mm Whatmann filter paper (Whatmann, UK) were placed in a 1.5ml microcentrifuge tube and 180ul of Animal Tissue Lysis (ATL) buffer was added followed by 85^0C incubation for 10 minutes and addition of 20ul proteinase K stock solution. The mixture was vortexed and incubated at 56^0C for 1 hour. Buffer AL was added to the sample, mixed thoroughly by vortexing and incubated at 70^0C for 10 minutes. Absolute

ethanol (200ul) was added to the sample and mixed thoroughly by vortexing. The mixture was then added to a QIAamp Mini spin column in a 2ml collection tube and centrifuged at 6000 xg (8000 rpm) for 1 minute. The spin column was placed in a clean 2 ml collection tube and the filtrate discarded alongside the tube. Buffer AW1 (500ul) was added and the mixture centrifuged at 6000 x g (8000 rpm) for 1 minute. The collection tube containing the filtrate was discarded. Buffer AW2 (500ul) was added to the spin column and centrifuged at full speed (20, 000 x g or 14, 000 rpm) for 3 minutes. The filtrate was again discarded and the spin column was placed in a 1.5 ml microcentrifuge tube. Parasite DNA was eluted in 150ul buffer AE, incubated at room temperature for 1 minute and centrifuged at 6, 000 x g (8 000 rpm) for I minute. The extracted DNA was stored at -20^{0}C and transported to The UK Medical Research Council Laboratories, The Gambia for subsequent molecular analyses.

3.7 MSP-1 AND MSP2 GENOTYPING AND ELECTROPHORESIS

Genotyping of Merozoite Surface Proteins 1 and 2- MSP-1 (block 2) and MSP2 (block 3) were carried out as previously described by Snounou and Beck (1998). The primer sequences are as shown in Appendix 2. PCR amplification was performed on thermal cycler (Techne, UK) in a final volume of 15µl. Cycling conditions the primary PCR of both MSP-1 and MSP-2 amplifications were as follows; Denaturation at 95^{0}C for 5 minutes, annealing at 58^{0}C for 2 minutes, 72^{0}C for 2 minutes; (95^{0}C for 1minute, 58^{0}C for 2 minutes, 72^{0}C for 2 minutes) x25 cycles; 58^{0}C for 2 minutes, and a final extension at 72^{0}C for 5 minutes. Two microlitres of primary PCR product was used as a DNA template in the secondary PCR which had similar concentrations to the primary PCR. The cycling conditions for the secondary PCR were as follows: Denaturation at 95°C for 10 minutes; (94°C for 30 seconds; annealing at 58°C for 30 seconds; 72°C for 1 minute) x40 cycles and a final extension at 72°C for 10 minutes. After

electrophoresis on 2% agarose gel using 0.5x TBE buffer at 100 volts, the PCR products were visualized by ultraviolet transillumination on gel documentation system (Upland, USA) after electrophoresis on 2% agarose gel using 0.5x TBE buffer at 100 volts.. Samples possessing only one genotype per allelic family were monoclonal while possession of multiple genotypes per family was described as polyclonality.

3.7.1 Statistical analysis of MSP genotyping

Data was analyzed using the SPSS software version 13. The relationships in the frequencies of the allelic families of MSP-1 and MSP2 loci between the study areas were tested using Chi-square.

3.8 MICROSATELLITE GENOTYPING

A two-round hemi-nested PCR was used to amplify 12 microsatellite loci from parasite DNA following previously described procedures and primers (Escalante *et al*., 1998). The loci included Poly α, TA42, TA81, TA87, TA1, TA109, TA40, 2490, ARAII, PfG377, PfPk2, and TA60 (Appendix 3). FAM, HEX and PET-labeled PCR products for different loci amplified from each isolate were pooled together with GeneScan™ 500 LIZ internal size standard (Applied Biosystems, Foster City, CA) for electrophoresis on an ABI 3130XL Genetic Analyzer. Peakscanner (Applied Biosystems) and GeneMarker (Softgenetics) softwares were used for normalization across runs and automatic determination of allele length and peak heights in samples containing multiple alleles per locus, minor alleles were scored when the minor peaks were ≥ 20% the height of the predominant allele in the isolate and with a relative fluorescent unit of at least 100.

3.9 MULTIPLICITY OF INFECTIONS

Multiple infections were defined when any of the loci contained multiple alleles. The Multiplicity of Infection (MOI) or number of genotypes per infection was calculated by dividing the total number of fragments detected in one antigenic marker by the number of samples positive for the same marker. Samples possessing only one genotype per allelic family were considered monoclonal while possession of multiple genotypes per family was described as polyclonality.

3.10 POPULATION GENETIC ANALYSES

3.10.1 Allele frequency

The allele frequencies, numbers of alleles per locus, allelic diversity within each population, and allele frequencies per locus per population were calculated using GENALEX 6 (Peakall and Smouse, 2006).

3.10.2 Allelic diversity

Allelic diversity was calculated for each of the microsatellite loci based on the allele frequencies, using the formula for expected heterozygosity,

$$H_E = [(n/n-1)\ (1-\sum p^2)] \quad \ldots\ldots\ldots\ldots\ldots \quad \textbf{Equation 3.3}$$

Where n is the number of isolates analyzed and p represents the frequency of each different allele at a locus. H_E provides an indication of the probability that two individuals will be different. It has a potential range from 0 (no allele diversity) to 1 (all sampled alleles are different).

3.10.3 Linkage disequilibrium

To understand the potential for multilocus haplotypes to spread through the populations, multilocus Linkage Disequilibrium (LD) was calculated for the entire population as a whole, and separately for each subpopulation using the standardized index of association, (I_A^S), using LIAN version 3.5 web interface (Haubold *et al.*, 1998) and the majority allele at each locus in each infection. This index was calculated as

$$(I_A^S) = (1/n - 1\ ((V_D / (V_E) - 1) \ldots\ldots\ldots\ldots \textbf{Equation 3.4}$$

Where V_E is the expected variance of n - the number of loci for which two individuals differ. The observed variance is given by V_D. To test whether the ratio of V_D/V_E was significantly greater than one, randomization test was carried out as previously described (Souza *et al.*, 1993; Gourdon, 1997).

3.10.4 Analysis of population differentiation

Between population and within population variance was determined with Analysis of Molecular Variance (AMOVA), (ΦPT), the analogue of Wright's Fst as it is flexible enough to accommodate different types of assumptions about the evolution of microsatellites (Excouffier e*t al*., 1992). $\Phi PT = 0$ was considered indicative of no genetic difference among populations. A distance between isolates from the different populations was estimated in GENALEX 6 which was also employed in implementing a principal coordinate analysis (PCoA) to determine population substructure. Population structure was visualized in an R-dot plot of coordinates 2 and 3.

3.11 NEXT GENERATION SEQUENCING OF *PLASMODIUM FALCIPARUM* GENOME

3.11.1 Leukocyte depletion using CF11 assembly

Venous blood collected from each participant in sterile ethylene diamine tetracetic acid (EDTA) – coated tubes was depleted of white blood cells by size selection and hydrophoresis using CF11 columns to reduce the presence of human DNA (Venkatesan *et al*., 2012). The CF11 columns were prepared by adding 2.0g of loosely packed cellulose powder to a syringe barrel. The plunger of the syringe was re-inserted and the column suspended over an uncapped 50ml tube after which 4ml of 1x Phosphate Buffer Saline (PBS) was pipetted onto the top of the column and allowed to flow through. Two ml of malaria-positive blood sample was applied to the top of the column and left to completely sip through. The filtered sample was centrifuged at 2400 rpm (1000x g) at room temperature for 6 minutes to pellet the cells. The supernatant was removed and discarded while the pellet was stored at -20^{0}C prior to DNA extraction. Thick and thin blood films were prepared before and after cellulose filtration to confirm the complete removal of leucocytes.

3.11.2 DNA extraction from leucocyte-depleted whole blood

Extraction of genomic DNA was carried out using Qiagen Mini Kit (Qiagen, UK) following procedures described by Snounou *et al*. (1998). Two hundred ul of whole blood was mixed with 20ul proteinase K. Buffer AL (200ul) was added to the mixture and pulse-vortexed for 15 seconds, incubated for 10 minutes at 56^{0}C and then centrifuged briefly. The mixture was applied to the QIAmp mini spin column (in a 2ml collection tube) without wetting the rim and centrifuged at 6000 x g (8000 rpm) for 1 minute. The spin column was carefully opened and 500ul of buffer AW1 was added and the mixture was centrifuged at 6000 x g (8000 rpm) for 1

minute. The process was repeated with buffer AW2; centrifuging at 20 000 x g (14 000 rpm) for 3 minutes after which the spin column was placed in a clean 1.5ml microcentrifuge tube and 200ul distilled water was added, incubated for 1 minute at 25^0C for elution and centrifuged at 6000 x g for 1 minute. The extracted DNA samples were stored at -20^0C for subsequent analysis. The quality as well as quantity of each DNA sample was determined using nanodrop (Thermo Fisher Scientific, USA).

3.11.3 Quantitation of parasite DNA post-depletion

Real-time PCR was carried out to determine the quantity of parasite DNA relative to human DNA contained in each sample using the procedures described by Veron *et al.* (2009). A master mix was prepared following the recipe described in Appendix 5. Twenty microlitres of the mix was dispensed to individual wells of a bar-coded plate and 5ul of template DNA was added to each well except in the no template control wells and loaded on BIO-RAD CFX96 Real-Time system. The real-time PCR conditions involve an initial activation step at 95^0C for 20 seconds followed by 40 cycles of denaturation at 94^0C for 3 seconds, annealing at 60^0C for 30 seconds and extension at 72^0C for 30 seconds. *Plasmodium* tRNA methionine (PgMET) gene was targeted in the parasite DNA while β tubulin gene was targeted in the human DNA. The primer sequences are as shown in Appendix 5.

3.11.4 Microbiome enrichment of parasite DNA

To further deplete human DNA extracted with the parasite DNA, NEBNext Microbiome DNA Enrichment Kit (New England Biolabs, UK) was used. Human genome is CG-rich and methylated. Methylated human DNA was removed from the DNA mixture by binding to the methyl-CpG binding domain of human MBD2-Fc protein (MBD2-Fc refers to the MBD domain

of the MBD2a protein fused to the Fc portion of the human antibody heavy chain). This stable complex selectively bound double-stranded methylated CpG-containing DNA. A sample of human (foetal lung fibroblast) IMR 90/*E. coli* (10:1 ratio) genomic DNA was included as a control. Enriched parasite DNA was loaded on a 0.8% agarose gel alongside an appropriate DNA marker (Lambda DNA-Hind III digest).

3.11.5 Reverse Transcriptase PCR detection of Human ribonuclease P

Detection of possible presence of human DNA in the final enriched samples was carried out using Taqman Ribonuclease P PCR kit (Applied Biosystems, UK). The reaction was done in one-step real time PCR. The first step was a Reverse Transcription (RT), during which the Rnase P was transcribed into cDNA. Afterwards, a thermostable DNA polymerase was used to amplify the specific gene fragments. Fluorescence was emitted and measured by the real time systems optical unit during the PCR. 1×10^7 copies/ml of human Rnase P was used as the positive control using 45^0C for 10 minutes (1 cycle), 95^0C for 15 minutes (1 cycle) and 40 cycles of 95^0C for 15 seconds and 60^0C. The reaction recipe is described in Appendix 6

3.11.6 Pooling of DNA, whole-genome isothermal amplification and enzymatic shearing

Concentrations of individual parasite samples were normalized to 500,000 parasites/ul in a final volume of 920ul of pooled DNA (Appendix 7). Whole genome amplification (WGA) of the DNA was carried out to non-specifically amplify pooled genomic DNA using Rapisome™ pWGA kit (Biohelix, England). The amplified DNA was sheared using NEBNext dsDNA fragmentase (New England Biolabs). To achieve the required fragment size of 200-1000bp, incubation was carried out at 37^0C for 20 minutes after which 5µl of 0.5M EDTA was added to stop the reaction.

3.11.7 DNA library preparation

To remove smaller DNA fragments of 150bp and below, GeneRead^{TM} size selection kit (Qiagen, UK) was used. Four volumes of buffer SB1 were added to 1 volume of the sheared DNA samples and mixed. The mixture was then added to a MinElute spin column and centrifuged for 1 minute. The flow-through was discarded and the MinElute spin column was placed back into the same tube. Ethanol was added to the MinElute column to wash after which buffer EB was added for elution. The ends of the sheared DNA strands were repaired by incubating for 20 minutes at 25^0C, followed by 10 minutes at 70^0C using the recipe described below in Appendix 7. Adapters were added to the end-repaired DNA and mixed thoroughly using the recipe described below and incubated for 10 minutes at 25^0C followed by 5 minutes at 72^0C (Appendix 8). GeneRead size selection kit (Life Technologies, United Kingdom) was used to select adapter-ligated DNA in the 280-320bp range for expected read lengths of 200bp. The selected size ranges were validated using an automated gelling system (QIAgen, Valencia, CA).

3.11.8 Dilution of Library

The quality of library DNA was assessed using the bioanalyzing package offered by QiAxcel advanced automated gelling system (Qiagen, United Kingdom). Library DNA was quantified using nanodrop (Thermo Fisher Scientific, United States of America) and library dilution was carried out as described below:

Library concentration = 31.32ng/ul

Average fragment size = 269bp

Molecular mass of a bp = 650g/mol/bp

Library molecular weight = 269bp x 650 g/mol/bp = 175012.5g/mol

Molar concentration = 31.32ng/ul / 175012.5g/mol

= 0.000179nmol/ul

= 0.179uM

=179nM

=179000pM

But 25ul of library was required for 26pM

Using $C_1V_1 = C_2V_2$,

where C1 = 26pM; V1 = 25ul; C2 = 179000pM; V2 = ?

V2 = 0.0036ul of library was used for one-touch (OT) 2 emulsion PCR.

3.11.9 Emulsion PCR and enrichment of library

The diluted library was vortexed and centrifuged and was mixed with 25ul of nuclease free water, 500ul of reagent mix, 300ul of PCR reagent B and 50ul of enzyme mix. The solution was vortexed and centrifuged prior to the addition of ion sphere particles. The amplification solution was loaded on the reaction filter assembly port and run on Ion OneTouch™ 2 instrument (Life Technologies, UK). The template-positive ion sphere particles were recovered after the run and enriched using the Ion OneTouch™ ES instrument (Life Technologies, UK). Melt-off solution was prepared by adding 280ul Tween solution to 40ul 1M NaOH. Thirteen microlitres of Dynabeads MyOne Streptavidin C1 beads contained in the Ion OneTouch™ ES kit (Life Technologies, UK) was transferred to a 1.5ml eppendorf tube and placed on a magnetic rack for two minutes after which the supernatant was discarded without disturbing the pellet of Dynabeads MyOne Streptavidin C1 beads. One hundred and thirty microlitres of MyOne beads wash solution was added to the streptavidin C1 beads before the tube was removed from the magnetic rack and vortexed. Melt-off solution, streptavidin and MyOne beads wash solution were mixed and placed in a magnetic rack and the supernatant containing the enriched ion sphere particles were collected for chip loading.

3.11.10 Chip loading and genome sequencing

Five microlitres of control ion sphere particles was added to the enriched ion sphere particles and centrifuged for two minutes at 15 500 x g. The supernatant was carefully removed and leaving 15ul of the ion sphere particles in the tube. Twelve microlitres of sequencing primer was added and the resulting mixture was pipetted up and down thoroughly to disrupt the pellet before being run in a thermal cycler at 95^0C for 2 minutes and then 37^0C for 2 minutes to anneal the sequencing primer. Three microliters of Ion PGM sequencing polymerase (Life Technologies, UK) was added to the annealed ion sphere particles, mixed and incubated at room temperature before loading on Ion 318^{TM} chip (Life Technologies, UK). The workflow is described below in Figure 3.2.

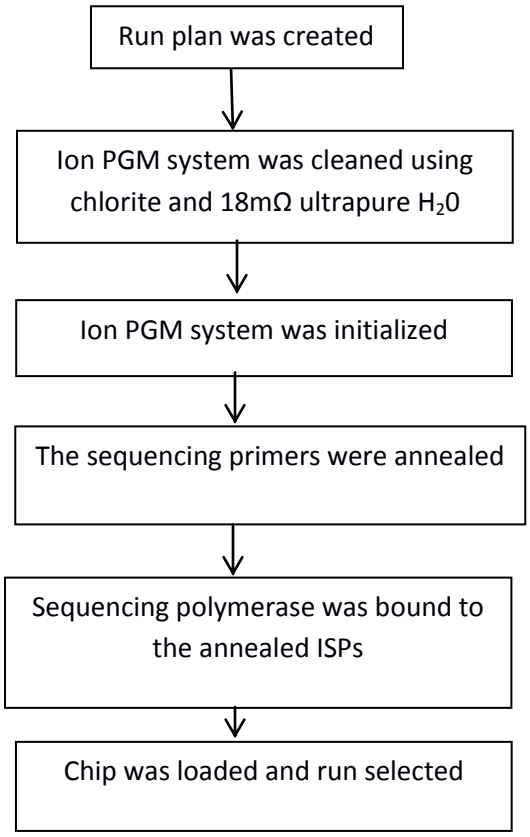

Figure 3.2 Workflow for pooled DNA sequencing

3.12 BIOINFORMATICS PIPELINE ANALYSIS AND DATA PROCESSING VIA COMMAND LINE TOOLS

Downstream analyses of genome sequence data were carried out in the command terminals of LINUX operating system (Ubuntu 14.04 LTS). Alignment of reads generated from the sequencing machine was done using mapping and assembly with quality (MAQ) tools (Figure 3.3). *P. falciparum* 3D7 genome, the fasta format of which was downloaded from PlasmoDB (www.PlasmoDB.org), was used as the reference sequence. The command lines used for analysis are described below:

maqfasta2bfa **Plasmo3D7.fasta Plasmo3D7.bfa**

The fastq format of the sequence reads was also converted to the binary fastq format:

maq fastq2bfq**PlasmoNIG**.fastq**output.bfq**

The binary fastq read sequence was aligned to the binary fasta reference genome:

maq match **output.map Plasmo3D7.bfa output.fq**

And the read alignment was viewed:

maq mapview **output.map**

The mapping assembly was built and consensus sequences and qualities were extracted:

maq assemble **consensus.cns Plasmo3D7.bfa output.map**2 > **assemble.log**

maq cns2fq **consensus.cns** > **cns.fq**

SAMTOOL was downloaded and unpacked:

tar –jxvf samtools-0.1.8.tar.bz2

To build samtools:

```
make samtools
```

Alignment results in SAM format:

```
maq assemble consensus.cns Plasmo3D7.bfa output.map2  > output.sam
```

Conversion of SAM file to BAM file:

```
samtools view -S -b output.sam > output.bam
```

Sorting of BAM file:

```
samtools sort output.bam sortedoutput.bam
```

Indexing the genome assemblies:

```
samtools faidx Plasmo3D7.fasta

Index sortedoutput.bam
```

Mpile up/generation of VCF file:

```
samtools mpileup -g -f Plasmo3D7.fasta sortedoutput.bam > output.bcf
```

SNP calling:

```
bcftools view -bvcg output.bcf > snp.bcf
```

SNP filtering:

```
bcftools view snp.bcf | vcfutils.pl varFilter - > snp.vcf
```

SNP annotation:

```
java –Xmx4g –jar snpEff.jar snp.vcf > snp.ann.vcf
```

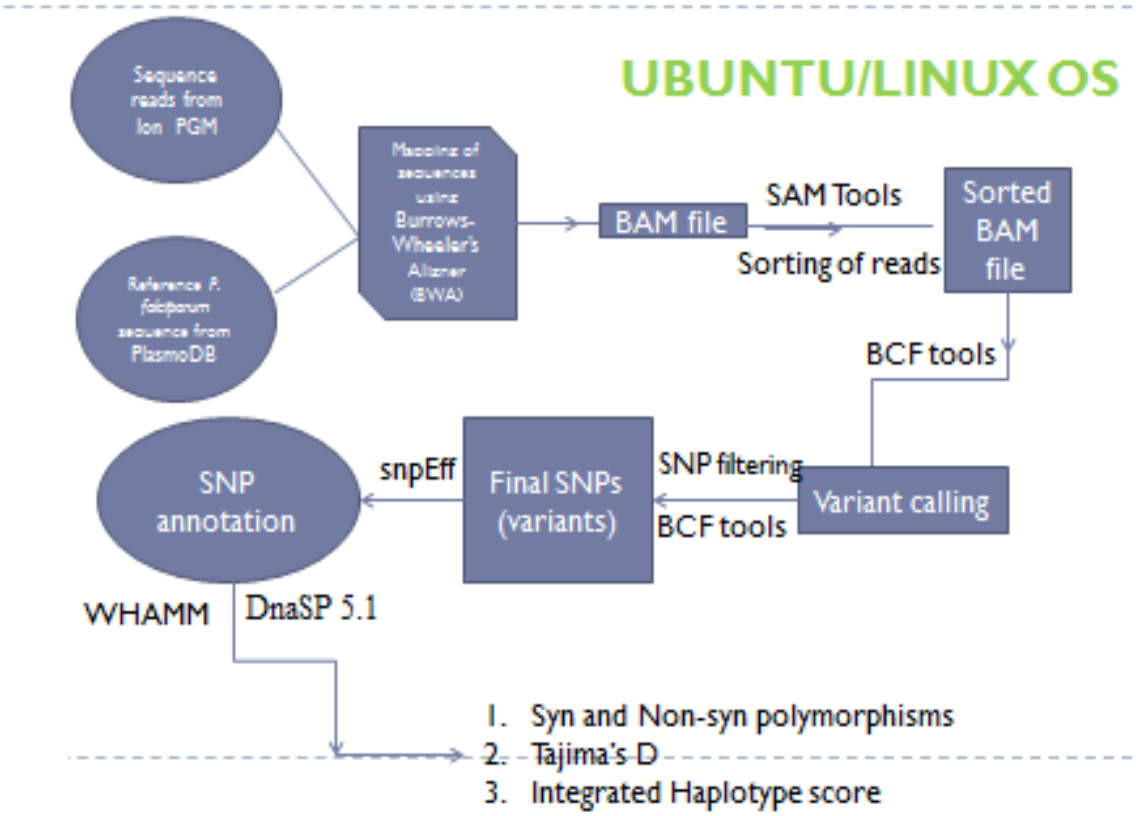

Figure 3.3 Algorithm for downstream computational analysis

3.13 STATISTICAL ANALYSIS OF SEQUENCE READS

3.13.1 Allele frequency determination

Allele frequencies in the population were determined for all SNPs, by analyzing all by analyzing all genotyped samples. Non-Reference (variant) Allele Frequency (NRAF) was computed as the proportion of genotyped samples whose allele was not the reference allele.

3.13.2 Determination of synonymous and nonsynonymous SNPs

All the typable SNPs defined in this study were in gene coding regions. Each SNP was classified as synonymous or nonsynonymous according to whether an amino acid change occurred when substituting the reference allele with the non-reference allele at that SNP in the 3D7 reference genome sequence, without any other changes. The reading frame and exon boundaries were determined from the PlasmoDB 13.0 annotation of the 3D7 genome (www.PlasmoDB.org).

3.13.3 Identifying genes under balancing and directional selections

An allele-frequency spectrum-based test, Tajima's D (Tajima, 1989), integrated haplotype score (Voight *et al*., 2006), and the ratio of nonsynonymous and synonymous polymorphisms (d_N/d_S) were applied to identify genes under selection. Tajima's D test takes into account the average pairwise nucleotide diversity between sequences (π) and the population nucleotide diversity parameter (Watterson's θ_w) expected under neutrality from the total number of segregating sites for a population at mutation-drift equilibrium (Tajima, 1989), with positive values (when $\pi > \theta_w$) indicating an excess of intermediate frequency polymorphisms and negative values indicating an excess of rare polymorphisms. DnaSP 5.1 (Rozas *et al*., 2003) was used to perform calculation of d_N/d_S ratios. Evidence of recent directional selection was obtained from the standardised haplotype score using WHAMM (Voight *et al*., 2006).

$$D = \frac{\pi - {}^{S}/_{\alpha 1}}{\sqrt{V}}$$ …… **Equation 3.5**

Where $\alpha_1 = \sum_{i=1}^{n-1} {}^{1}/_{i}$ i = index of summation and V = variance bêtween π and ${}^{S}/_{\alpha 1}$. DnaSP 5.1 (Rozas *et al*., 2006) was used to perform calculation of Tajima's D ratios. Evidence of recent directional selection was obtained from the standardised haplotype score using WHAMM (Voight *et al*., 2006).

$$iHS = ln\left(\frac{iHH_A}{iHH_D}\right)$$ …… **Equation 3.6**

Where, iHH_A is the integrated extended haplotype homozygosity for the ancestral allele and iHH_D is the integrated extended haplotype homozygosity for the derived allele.

Extreme positive iHS scores (>2.5) suggested selective sweeps along the ancestral allele signaling strong directional selection around that site.

CHAPTER FOUR

4.0 RESULTS

4.1 DEMOGRAPHIC PROFILES

The malaria parasites identified in all the study areas were *Plasmodium falciparum* species. Two thousand six hundred and seven participants were screened for uncomplicated malaria out of which 834 were confirmed *P. falciparum*-positive by microscopy. Male participants constituted 78.4% of particpants presenting with uncomplicated *falciparum* malaria in LEK. However, more females presented with symptoms of in AMK and BDG than the male participants. The proportion of patients that turned out to be positive for malaria was in the following order; AMK > BDG > LEK (Table 4.1).

4.2 PARASITE CLONAL DISTRIBUTION

Three hundred *P. falciparum* isolates were randomly selected and analysed for MSP-1 and MSP2 (100 from each study area). The individual size ranges for the amplified DNA fragments were 160-225bp, 130-220bp, 160bp, 400-700bp, and 470-520bp for K1, MAD20, RO33, FC27 and 3D7 respectively (Plates 1 – 5). All the three families of MSP-1 (K1, MAD20 and RO33) and two of MSP-2 (FC27 and 3D7) were observed among the isolates in each population. The frequency of isolates with K1 family was 66.7% in the overall population (i.e., 70/100 in LEK, 50/100 in BDG and 80/100 in AMK). The frequency of isolates with MAD20 was 56.7% in the overall population. RO33 proportions were 50/100 in LEK, 40/100 in BDG and 84/100 in AMK, and the family was observed to be monomorphic. The observed proportions indicating the allelic variants of families of MSP-1 and MSP-2 and clonal distribution in the study areas are presented in Table 4.2. Combinations of MSP-1 families observed were K1+MAD20, K1+R033 and

K1+MAD20+R033. None of the isolates had MAD20+R033 infections (Figure 4.1). The proportion of trimorphic infections (K1+MAD20+R033) was highest (90/300). Five percent of the parasite isolates had K1+MAD20 infections while 10% (30/300) of the isolates had K1+R033 families. For MSP-2, dimorphic infections with both 3D7 and FC27 allele types were detected among the isolates. The frequency of samples possessing FC27 type (66.7%) was found to be higher than the samples with only 3D7 family (54.7%). The MSP-2 infections with both allelic types were identified in 40% of the parasite isolates (Figure 4.2).

4.3 MULTIPLICITY OF INFECTIONS

Highest and lowest mean multiplicity of infections (MOI) with varying clones of *P. falciparum* were recorded in AMK (1.76) and BDG (1.29) respectively (Table 4.3) although the difference across the populations was not significant ($P = 0.637$). In Aramoko, samples from participants aged 2-4 years, 5-9 years, 10-17 years and 18-76 years had mean MOIs of 2.13, 1.83, 1.52 and 1.51 respectively. In Lekki, samples from participants aged 2-4 years, 5-9 years, 10-17 years and 18-76 years had mean MOIs of 1.86, 1.23, 1.31, 1.08 and 1.37 respectively. The mean MOIs for participants 2-4 years, 5-9 years, 10-17 years and 18-76 years in Badagry were 1.42, 1.01, 1.41 and 1.32 respectively (Figure 4.3). When the study populations were taken together, mean MOI was observed to be highest in individuals between 2-4 years (1.80) and lowest in individuals above 18 years. There was a significant difference in MOI between the 2-4 years age group and the 5-9 years age group as well as between the 5-9 years group and the 10-17 years group ($P < 0.01$). However, there was no significant difference in the MOI values between the 10-17 years age group and the 18-76 years age group ($P = 0.655$).

4.4 ALLELIC FREQUENCY

A total of 196 isolates of *P. falciparum* infections only were genotyped. Of the 12 microsatellite loci genotyped, 2 (TA87 and TA1) gave less efficient PCR amplification and were therefore excluded from subsequent analyses. The allelic frequencies at each of the ten loci in each of the three parasite populations are presented in Figure 4.4a to 4.4j. The overall number of alleles per locus observed in the study areas ranged from 8 (for locus 2490) to 27 (for locus TA81).

Table 4.1: Profile and prevalence of malaria among study participants

Profile	STUDY AREAS		
	Lekki (LEK)	Aramoko-Ekiti (AMK)	Badagry (BDG)
Participants screened for uncomplicated malaria	564	1309	784
Microscopically confirmed cases of *P.falciparum* infection	125	504	205
Prevalence (%)	22.1	38.5	26.1
Male	52	124	68
Female	73	380	137
Age range (mean) in years	2-37 (20.5±3.4)	2-76 (34.4±10.5)	2-54 (28.7±8.6)

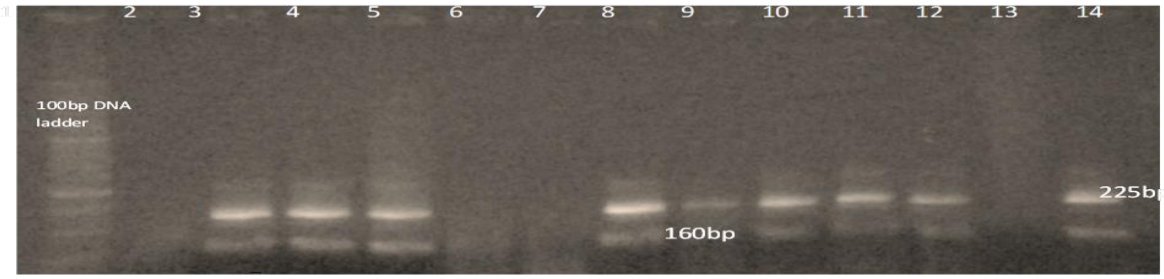

Plate 4.1 Molecular identification of MSP-1/K1 clones

Lane 1: 100bp DNA ladder

Lane 2: NTC

Lanes 3,4,5,8,9,10,11,12,13,15: Infections with two K1 clones

Lanes 6,7,14: Infections with no K1 clones

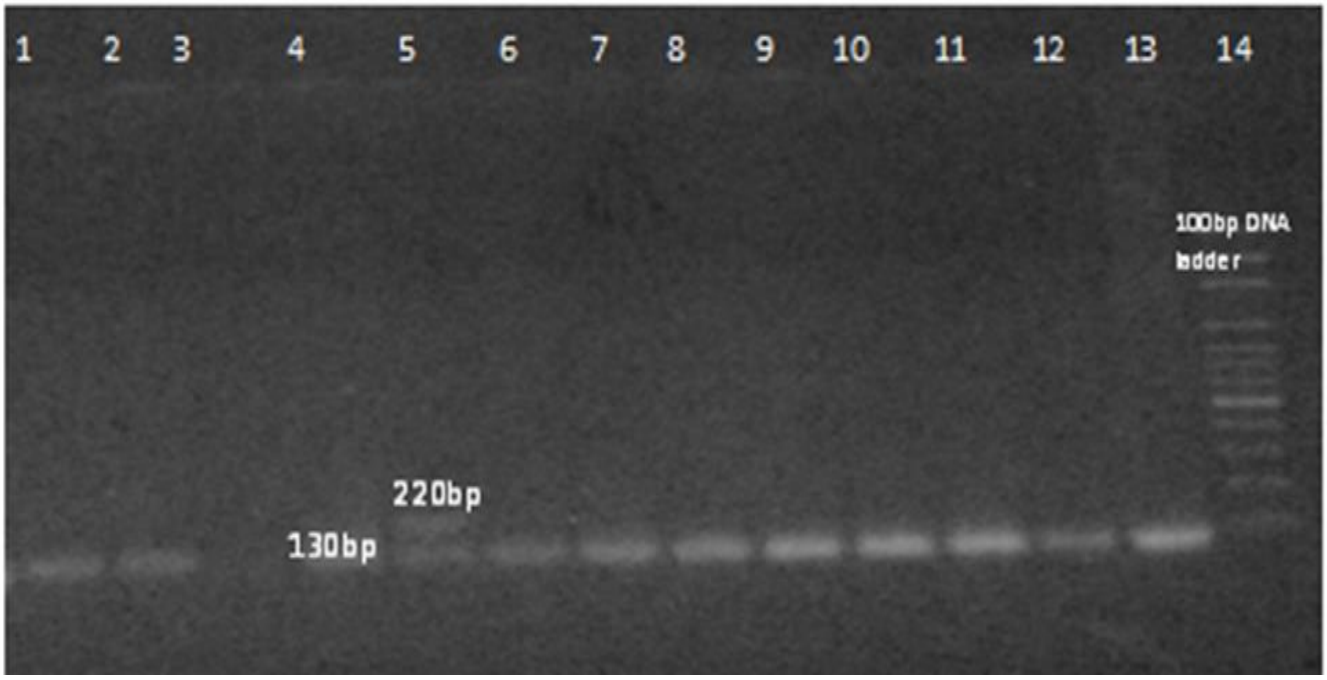

Plate 4.2: Molecular identification of MSP-1/MAD20 clones

Lanes1,2,4,5-13: Infections with single MAD20 clones

Lane 3: Infection with two MAD20 clones

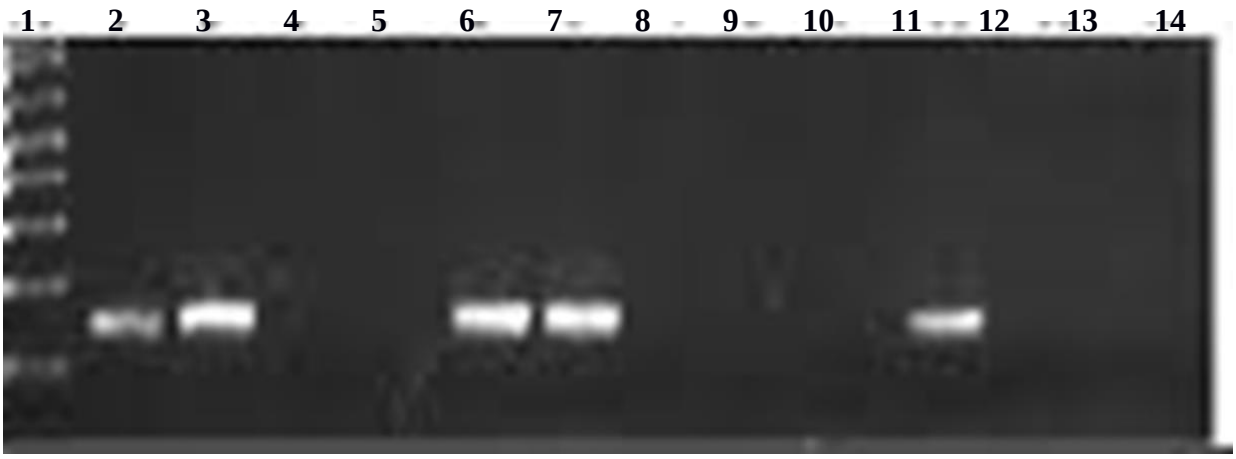

Plate 4.3: Molecular identification of MSP-1/RO33 clones

Lane 1: 100bp DNA ladder

Lane 2: RO33 Positive control

Lanes 4,5,8,9,10 and 13: Infections with no RO33 family

Lanes 2,3,6,7 and 11: Infections with RO33 family (160bp)

Lane 14: No template control

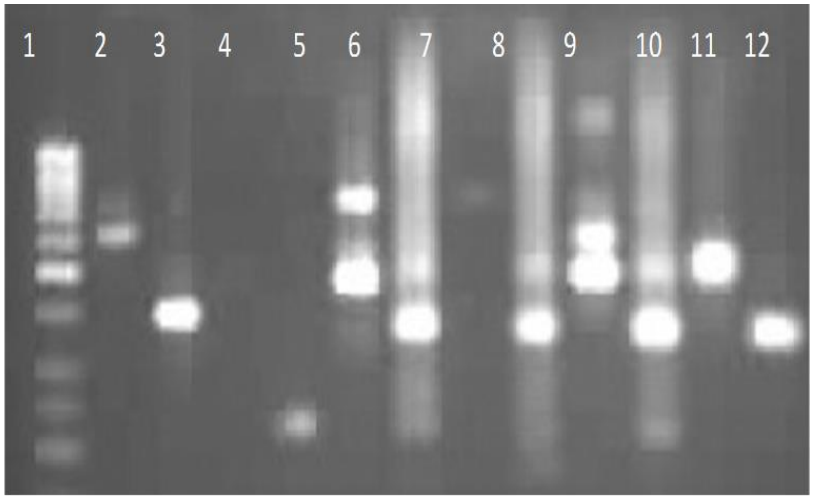

Plate 4.4: Molecular identification of MSP2/FC27 clones

Lane 1: 100bp DNA ladder

Lane 2: 610bp FC27 clone

Lanes 3 and 13: 400bp FC27 clones only

Lanes 7 and 11: 400bp and 500bp clones

Lane 6: 500bp and 700bp clones

Lane 10: 500bp and 600bp clones

Lane 12: 500bp only

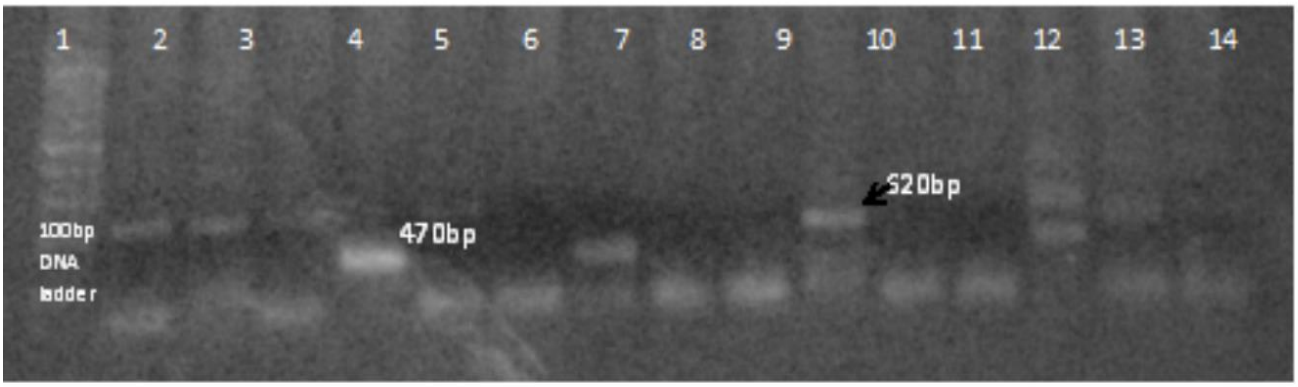

Plate 4.5: Molecular identification of MSP2/3D7 clones

Lane 1: 100bp DNA ladder

Lanes 2,3,4 and 15: 520bp 3D7 clones only

Lanes 12 and 15: 470bp and 520bp 3D7 clones

Table 4.2: Allelic variants and clonal diversity of parasite isolates from the study areas

Markers		Lekki	Badagry	Aramoko
MSP-1				
K1	n(%)	35 (70)	25 (50)	40 (80)
	Genotypes	2	3	3
MAD20	n(%)	25 (50)	25 (50)	35 (70)
	Genotypes	2	2	3
R033	n(%)	25 (50)	20 (40)	42 (84)
	Genotypes	1	1	1
MSP2				
FC27	n(%)	47 (94)	15 (30)	38 (76)
	Genotypes	3	1	4
3D7	n(%)	35 (70)	20 (40)	27 (54)
	Genotypes	2	2	3
MOI		**1.37**	**1.29**	**1.75**

MSP = Merozoite Surface Protein; MOI = Multiplicity of Infection

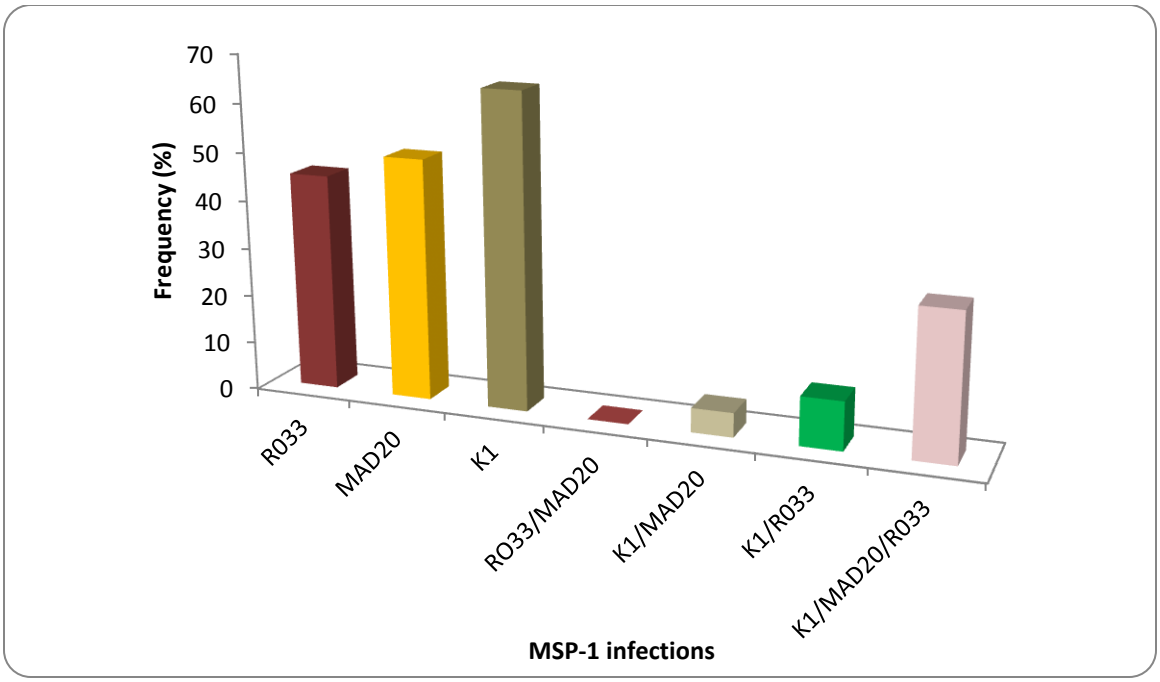

Figure 4.1: Frequency of MSP-1 families

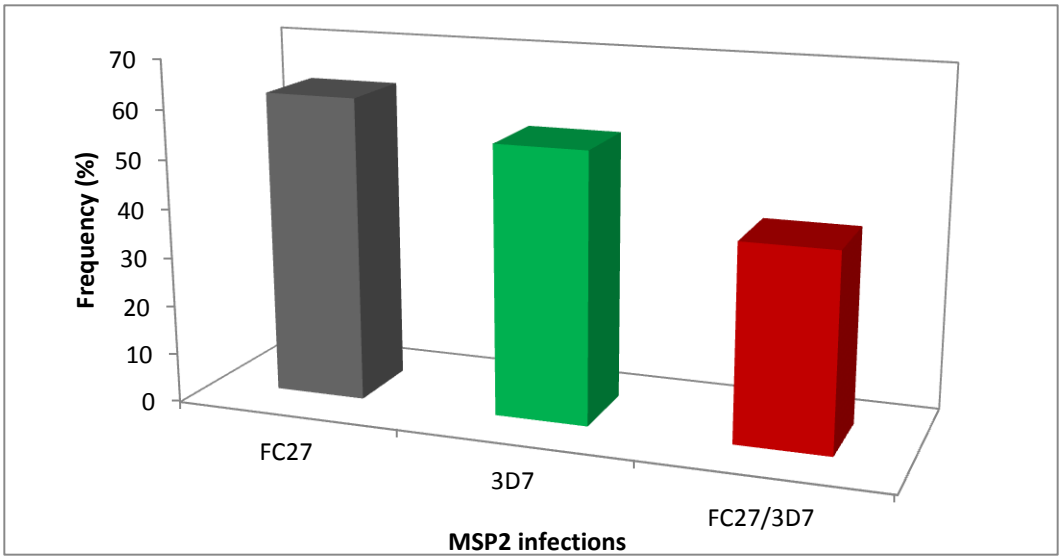

Figure 4.2: Frequency of MSP2 families

Table 4.3: Multiplicity of *Plasmodium falciparum* infections in the study areas

Locus	Lekki	Badagry	Aramoko
Poly α	1.60	1.59	1.65
PFPK2	2.00	2.44	2.40
TA81	1.89	2.14	1.96
ARA II	3.13	2.75	2.50
TA40	1.36	1.63	1.43
TA42	1.75	1.50	1.27
2490	1.11	1.33	1.00
TA60	1.18	1.44	1.33
TA109	2.08	1.59	1.47
PFG377	1.46	1.47	1.42
Mean MOI	**1.76**	**1.79**	**1.64**

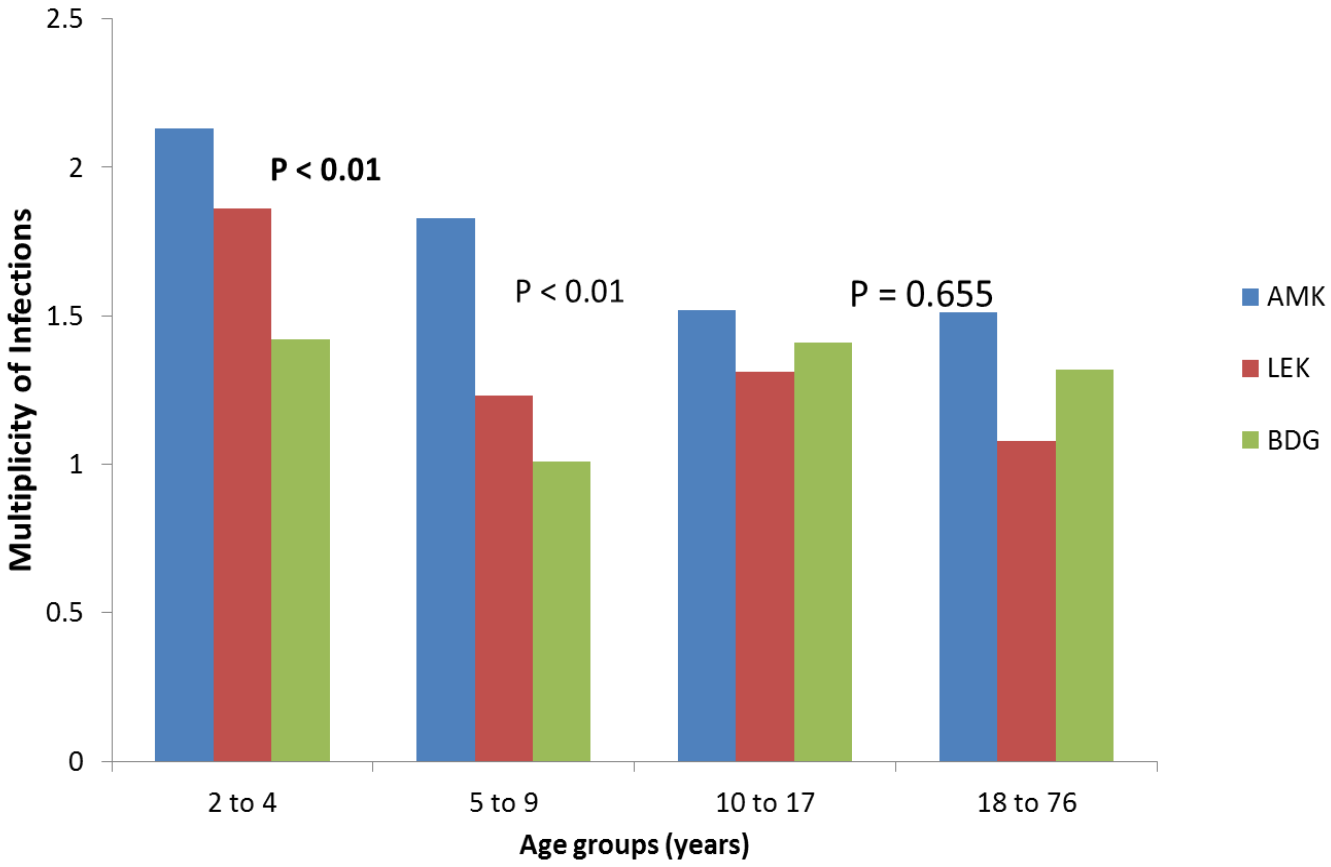

Figure 4.3: Multiplicity of infections by age group showing mean MOI decreased with increasing age

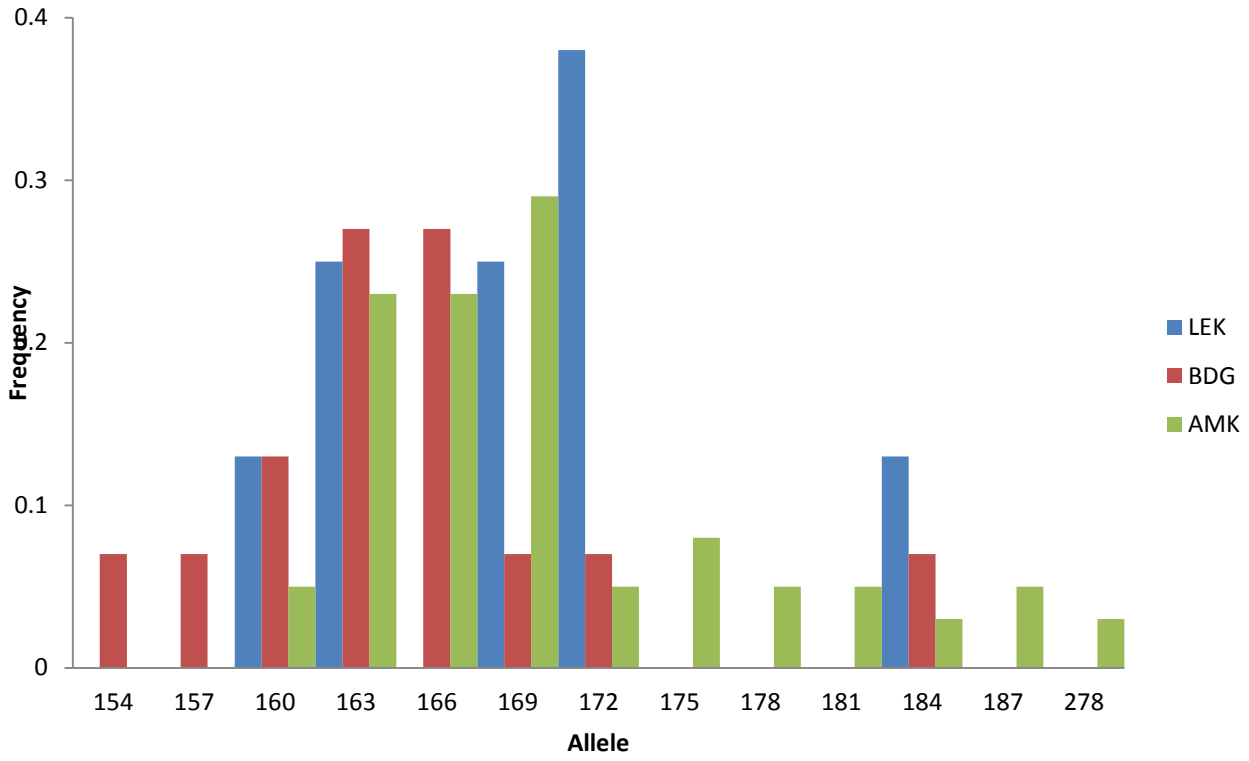

Figure 4.4a: Frequency distribution of the allele lengths (bp) at PfPK2 locus in the three *Plasmodium falciparum* populations

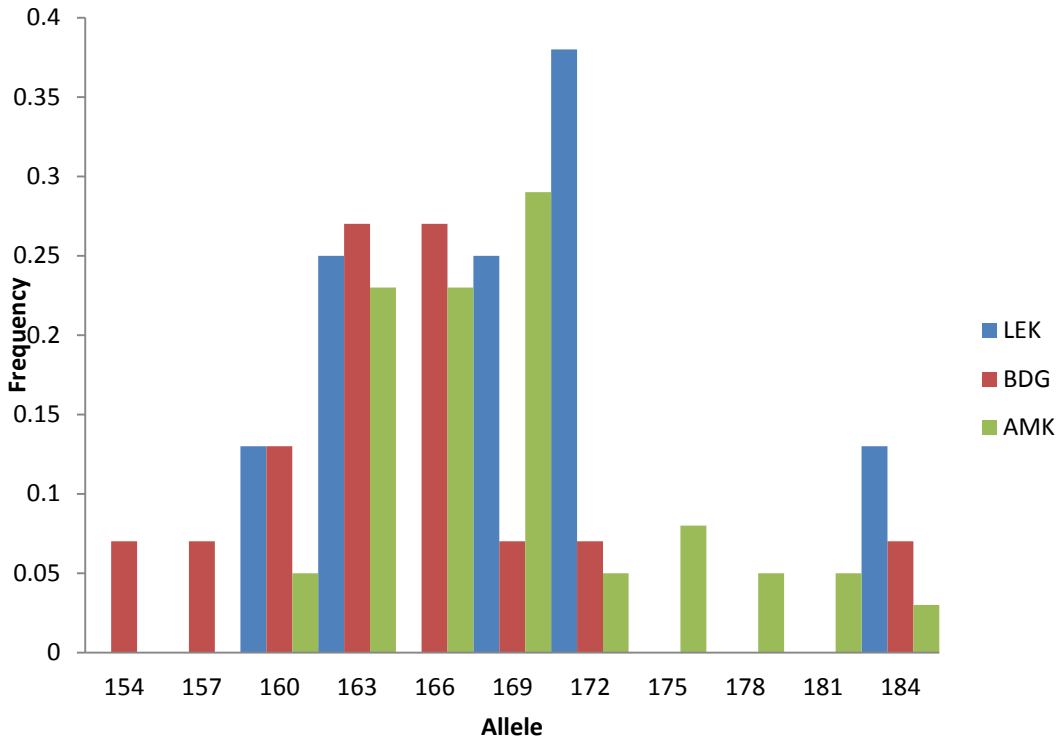

Figure 4.4b: Frequency distribution of the allele lengths (bp) at Poly α locus in the three *Plasmodium falciparum* populations

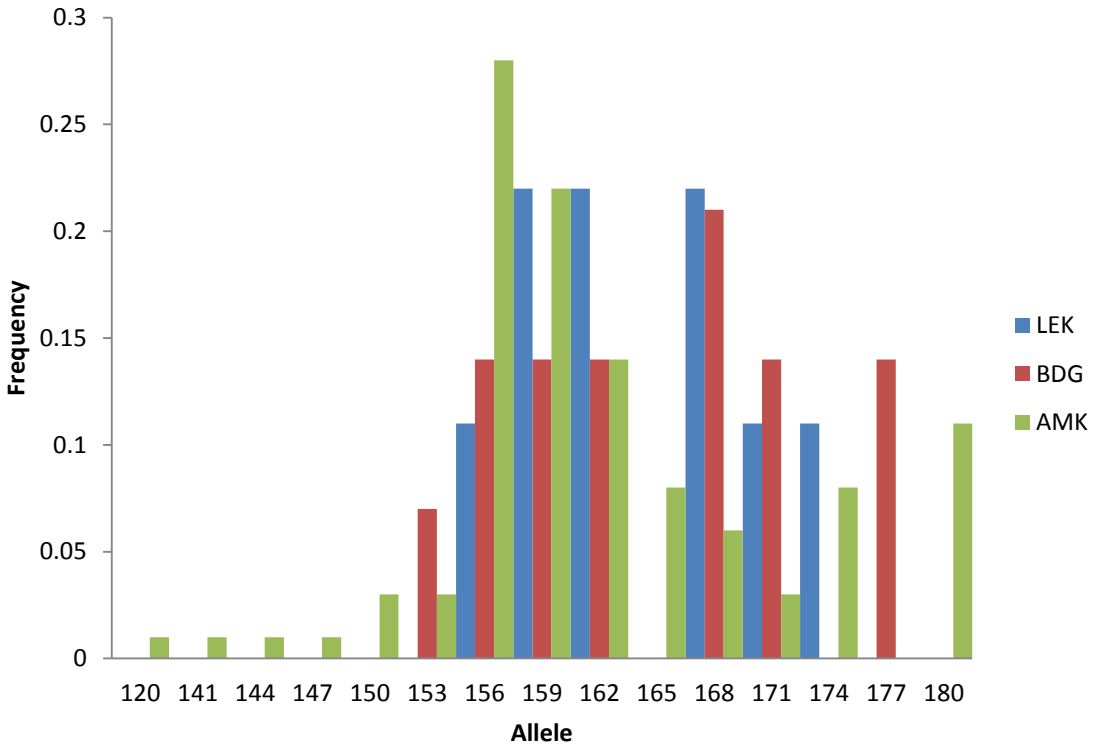

Figure 4.4c: Frequency distribution of the allele lengths (bp) at TA81 locus in the three *Plasmodium falciparum* populations

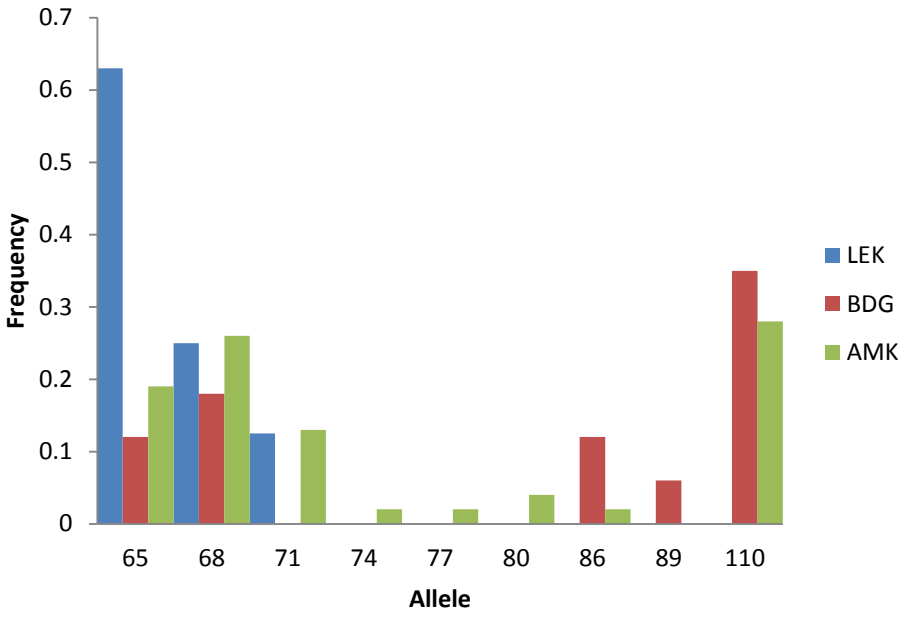

Figure 4.4d: Frequency distribution of the allele lengths (bp) at ARA II locus in the three *Plasmodium falciparum* populations

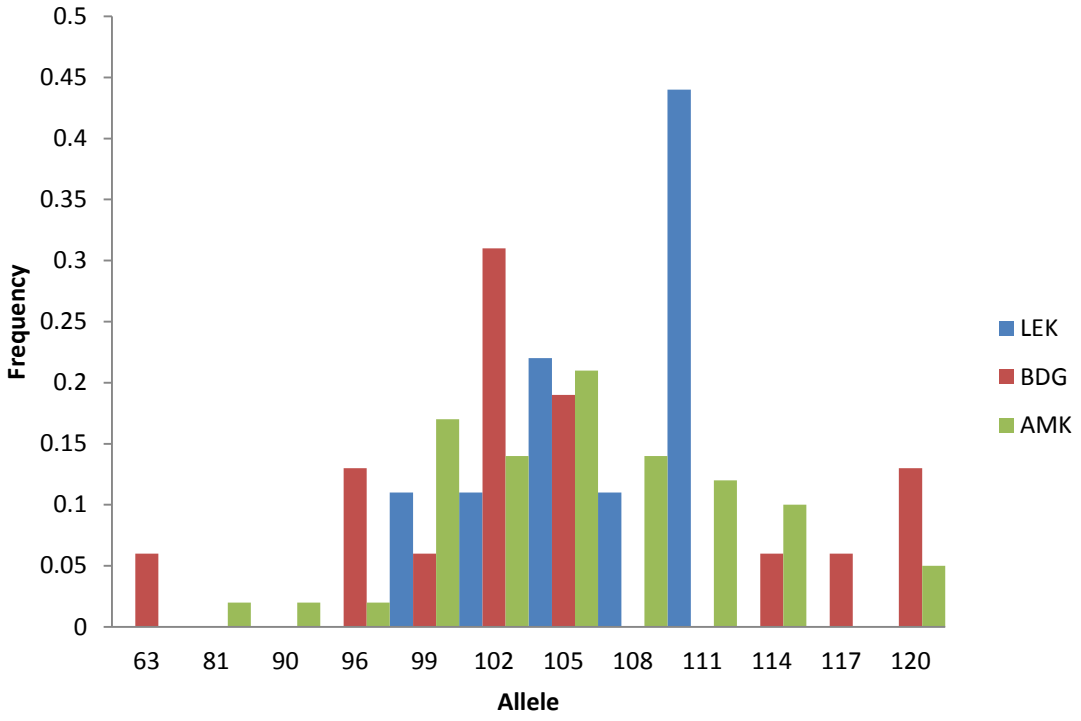

Figure 4.4e: Frequency distribution of the allele lengths (bp) at TA 40 locus in the three *Plasmodium falciparum* populations

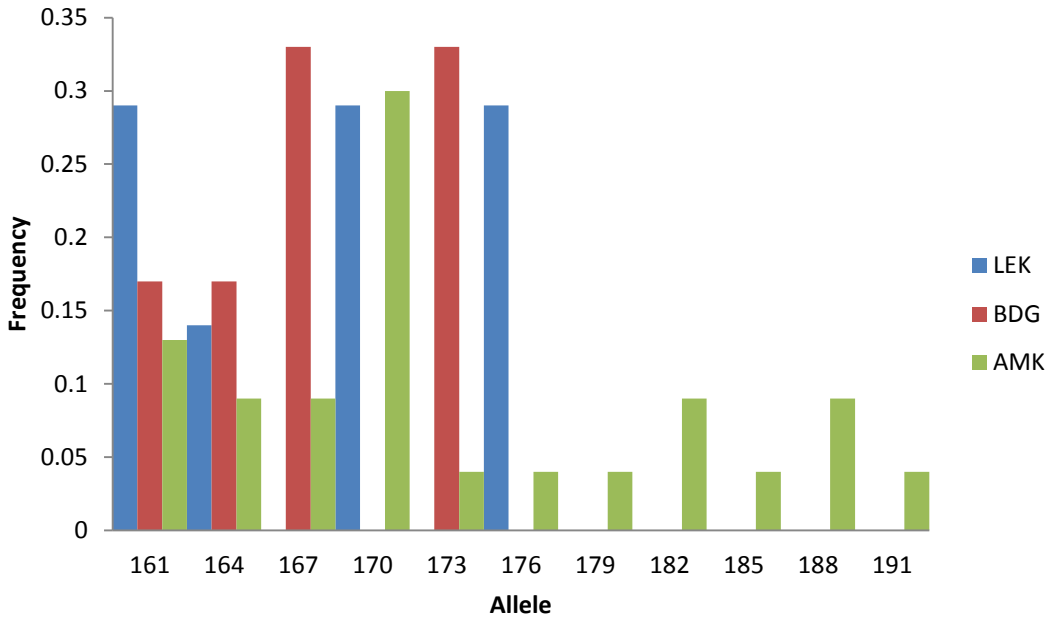

Figure 4.4f: Frequency distribution of the allele lengths (bp) at TA 42 locus in the three *Plasmodium falciparum* populations

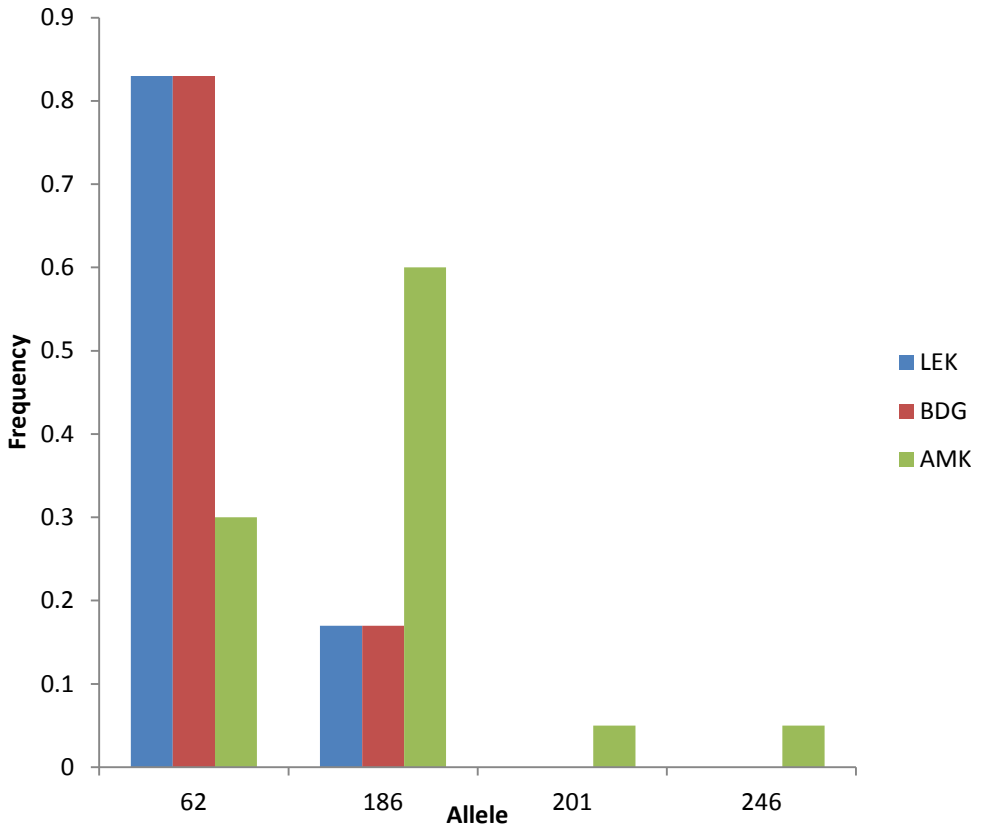

Figure 4.4g: Frequency distribution of the allele lengths (bp) at 2490 locus in the three *Plasmodium falciparum* populations

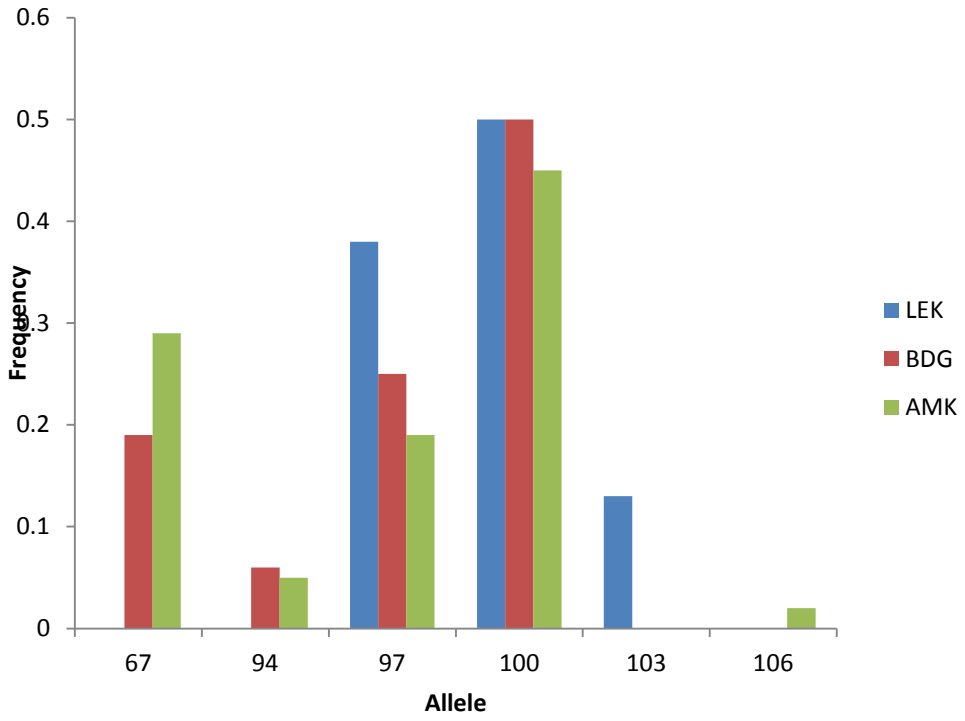

Figure 4.4h: Frequency distribution of the allele lengths (bp) at TA 60 locus in the three *Plasmodium falciparum* populations

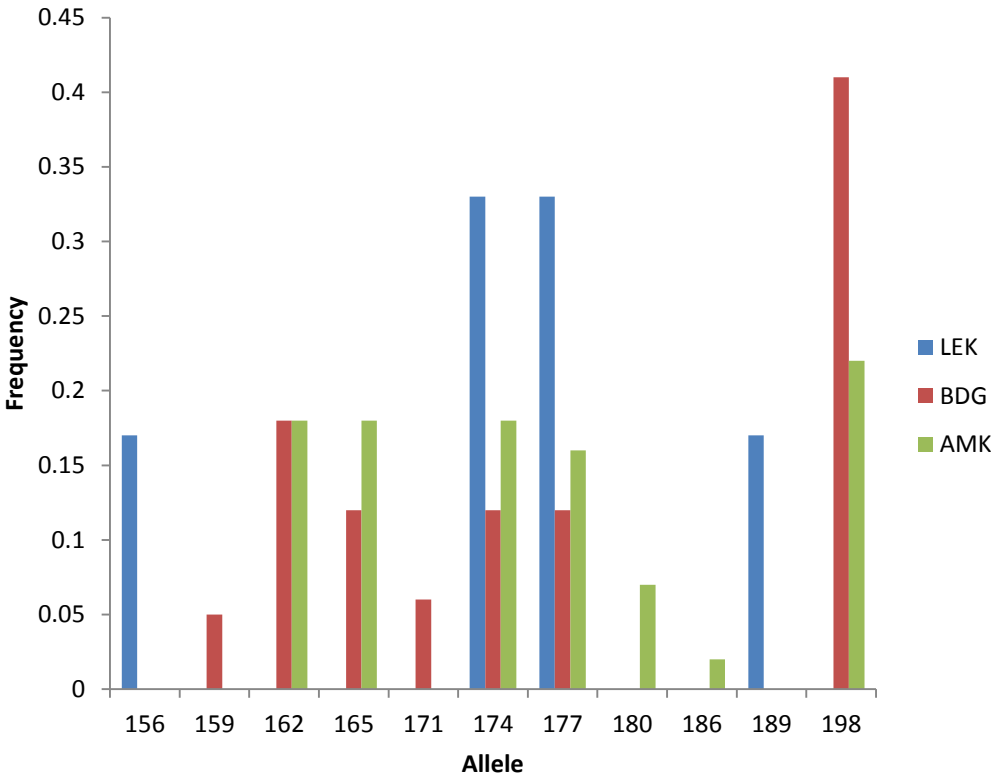

Figure 4.4i: Frequency distribution of the allele lengths (bp) at TA 109 locus in the three *Plasmodium falciparum* populations

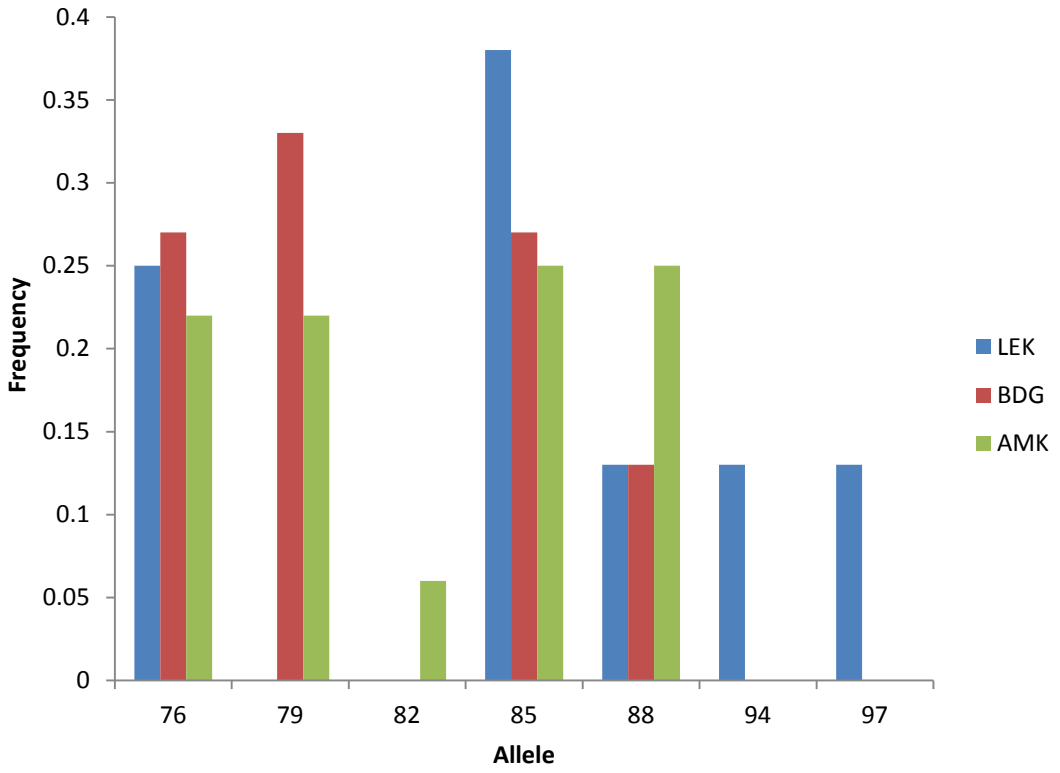

Figure 4.4j: Frequency distribution of the allele lengths (bp) at PfG377 locus in the three *Plasmodium falciparum* populations

4.5 ALLELIC DIVERSITY

Allelic diversity values were similar across the populations, with mean H_E values across all loci between 0.65 (for LEK) and 0.79 (for AMK) (Table 4.4). Mann-Whitney U-test result showed no significant difference in the mean H_E values between LEK and BDG as well as BDG and AMK at $P < 0.01$. However, the difference in the H_E values between LEK and AMK was significant ($P = 0.01$). Although the mean number of genotypes detected per isolate was highest in AMK (Table 4.5), Kruskal-Wallis test ($P=0.368$) showed no substantial difference in the mean number of genotypes in the three parasite populations.

4.6 POPULATION DIFFERENTIATION

Forty-three isolates (22%) had complete genotype data for all loci from which analysis of multilocus haplotypes was examined. Comparisons of populations using analysis of molecular variance (AMOVA) showed that genetic differentiation was low with $\Phi PT = 0.017$ ($P = 0.772$). Pairwise genetic distances between LEK and BDG, LEK and AMK and BDG and AMK parasite populations, calculated as Nei unbiased genetic distance (*u*D), were 0.164, 0.175 and 0.074 respectively. The relationship between genetic distance and the natural log of the geographical distance for each pair of parasite populations studied is presented in Figure 4.5. Principal coordinates analysis (PCoA) showed two distinct clusters of parasites not defined by the origins of individual population (Figure 4.6). AMOVA also indicated that almost all the genetic variations among parasites (99%) were contained within populations.

4.7 LINKAGE DISEQUILIBRIUM

Analysis of multilocus LD showed no significant index of association in all the parasite populations (Table 4.6).

Table 4.4: Allelic diversity (H_E) of microsatellite loci from the three parasite populations

Locus	Lekki	Badagry	Aramoko
Poly α	0.66	0.87	0.84
PfPK2	0.73	0.83	0.80
TA81	0.85	0.89	0.83
ARA II	0.55	0.81	0.82
TA40	0.75	0.84	0.91
TA42	0.77	0.77	0.87
2490	0.30	0.30	0.56
TA60	0.62	0.73	0.68
TA109	0.49	0.75	0.83
PfG377	0.78	0.75	0.79
Mean	**0.65**	**0.75**	**0.79**

Table 4.5: Assessment of the genotypes of each isolate

Locality	Isolates with given number of genotypes				Mean number of genotypes
	1	2	3	4	
Lekki	4	6	0	0	1.60
Badagry	9	5	1	1	1.59
Aramoko	18	18	5	0	1.66

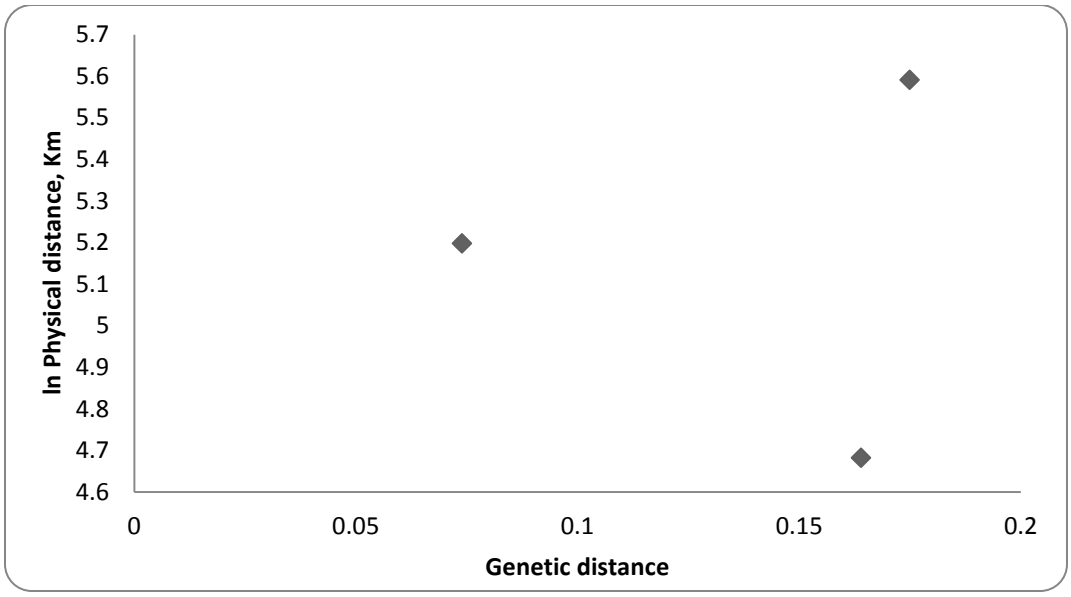

Figure 4.5: Relationship between geographic and genetic distances for each pair of parasite populations studied Genetic distance (y-axis) was determined using GENALEX 6.0 for each pair of populations separated by distance in kilometers (plotted on the x-axis in natural log scale).

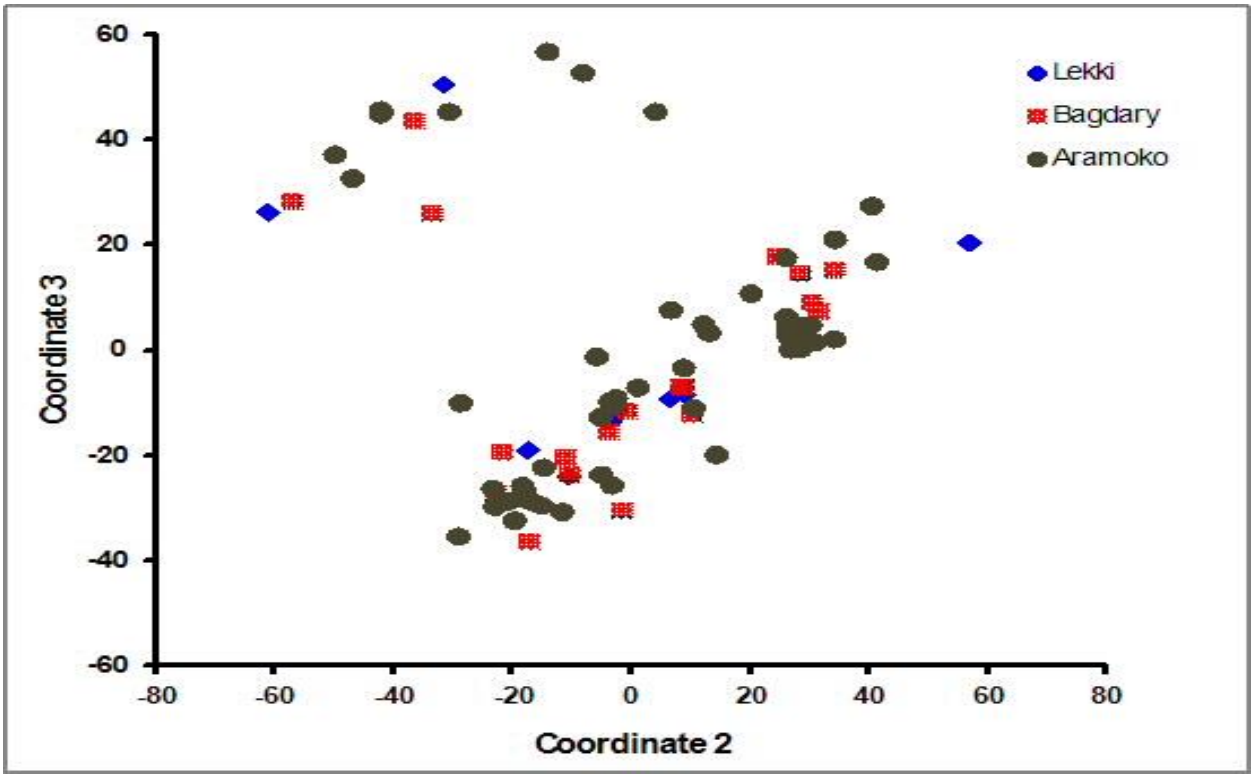

Figure 4.6: Population structure of *Plasmodium falciparum* isolates from southwestern Nigeria. Principal coordinate analysis (PCoA) from allelic variance at 10 haploid microsatellite loci. (Coordinates 2 and 3 in the PCoA show limited sub-structuring of *P. falciparum* isolates into two clusters not determined by site)

Table 4.6: Linkage disequilibrium analysis for each *Plasmodium falciparum* population

Population	V_D	V_E	I_A^S
Lekki	12.174	1.971	0.021
Badagry	7.484	1.657	0.032
Aramoko	4.318	1.544	0.062

V_E is the expected variance of n - the number of loci for which two individuals differ. V_D = the observed variance.

I_A^S =Linkage disequilibrium

4.8 NEXT GENERATION SEQUENCING

4.8.1 Quantitation of genomic DNA

Having established limited genetic differentiation between *P. falciparum* populations from southwestern Nigeria, one hundred falciparum-positive blood samples that had been leucocyte-depleted were subjected to genome sequencing. There were at least 5000 parasites/µl in the 100 isolates analysed. Normalisation to 500, 000 parasites/µl required estimated amounts of 4µl, 10µl and 20µl of representative samples in a final DNA volume of 920µl (Appendix 9).

4.8.2 Post-enrichment validation

Microbiome enrichment of the parasite DNA showed a further depletion of the host DNA in a control sample that had not been previously filtered with CF11 cellulose (Plate 4.6). The cycle threshold values before and after enrichment are indicated in Figures 4.7. The whole genome amplification picture is shown in Plate 4.7. Enzymatic shearing of DNA and subsequent size selection produced fragments between 179 and 466bp

Plate 4.6: Post-enrichment validation of parasite DNA

Lane 1: Lambda ladder
Lane 2: Unenriched IM-90/*E.coli* control
Lane 3: Unenriched pooled samples 1
Lane 4: Supernatant fraction from 1
Lane 5: Final enriched pooled samples 1
Lane6: Unenriched pooled samples 2
Lane 7: Supernatant fraction 2
Lane 8: Final enriched pooled samples from 2

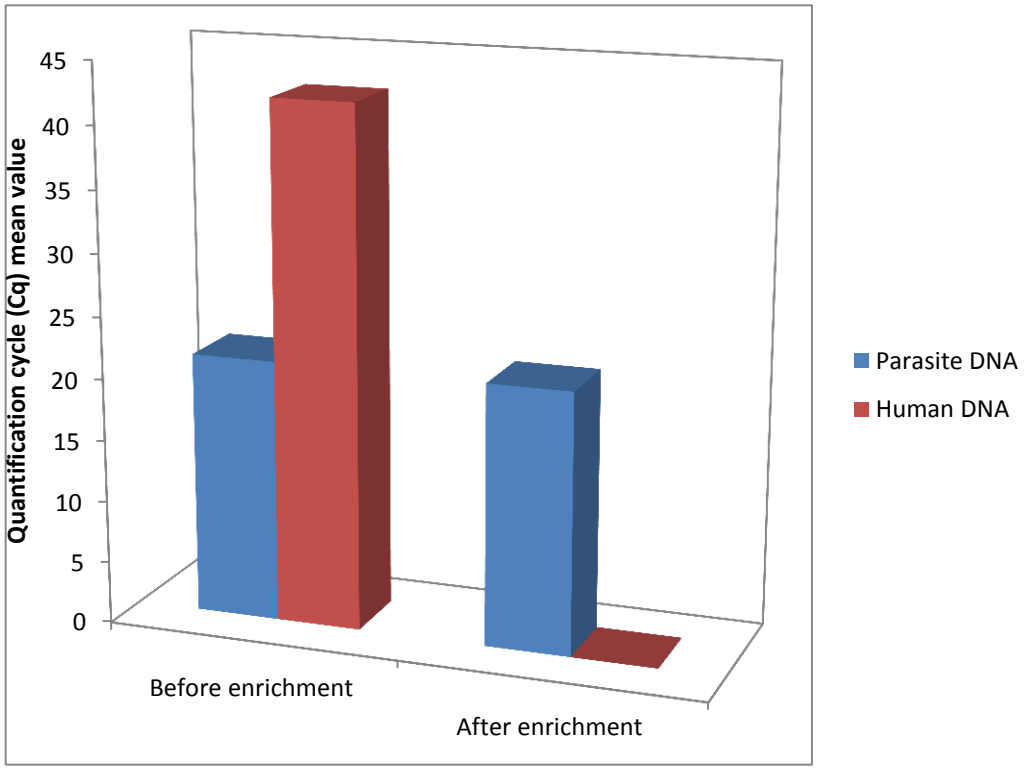

Figure 4.7: Cycle threshold values of parasite and human DNA in unenriched pool

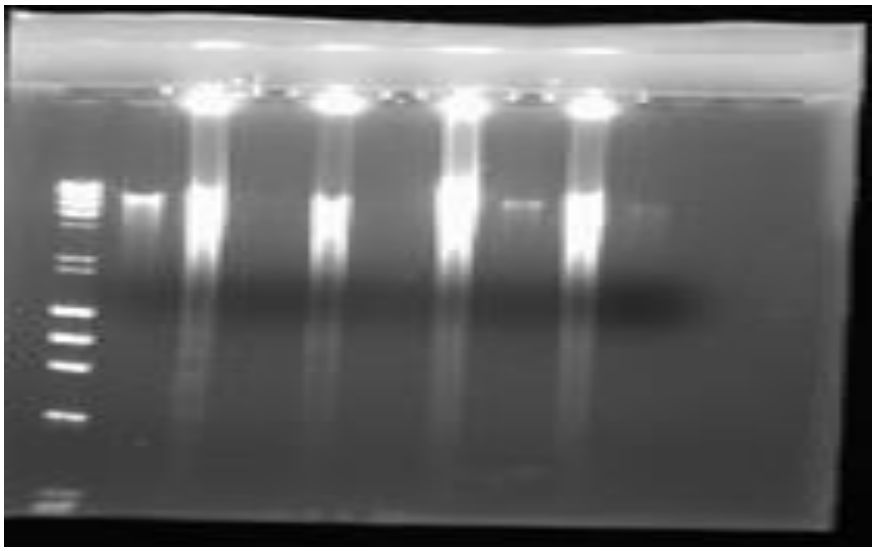

Plate 4.7: Whole genome isothermal amplification of enriched DNA

Lane 1: Lambda DNA ladder
Lane 2: Control DNA
Lane 3: Amplified unenriched pool 1
Lane 4: Unenriched unamplified pool 1
Lane 5: Amplified enriched pool 1
Lane 6: Unamplified enriched pool 1
Lane 7: Amplified pool 2
Lane 8: Unenriched pool 2
Lane 9: Amplified enriched pool 2
Lane 10: Unamplified enriched pool 2

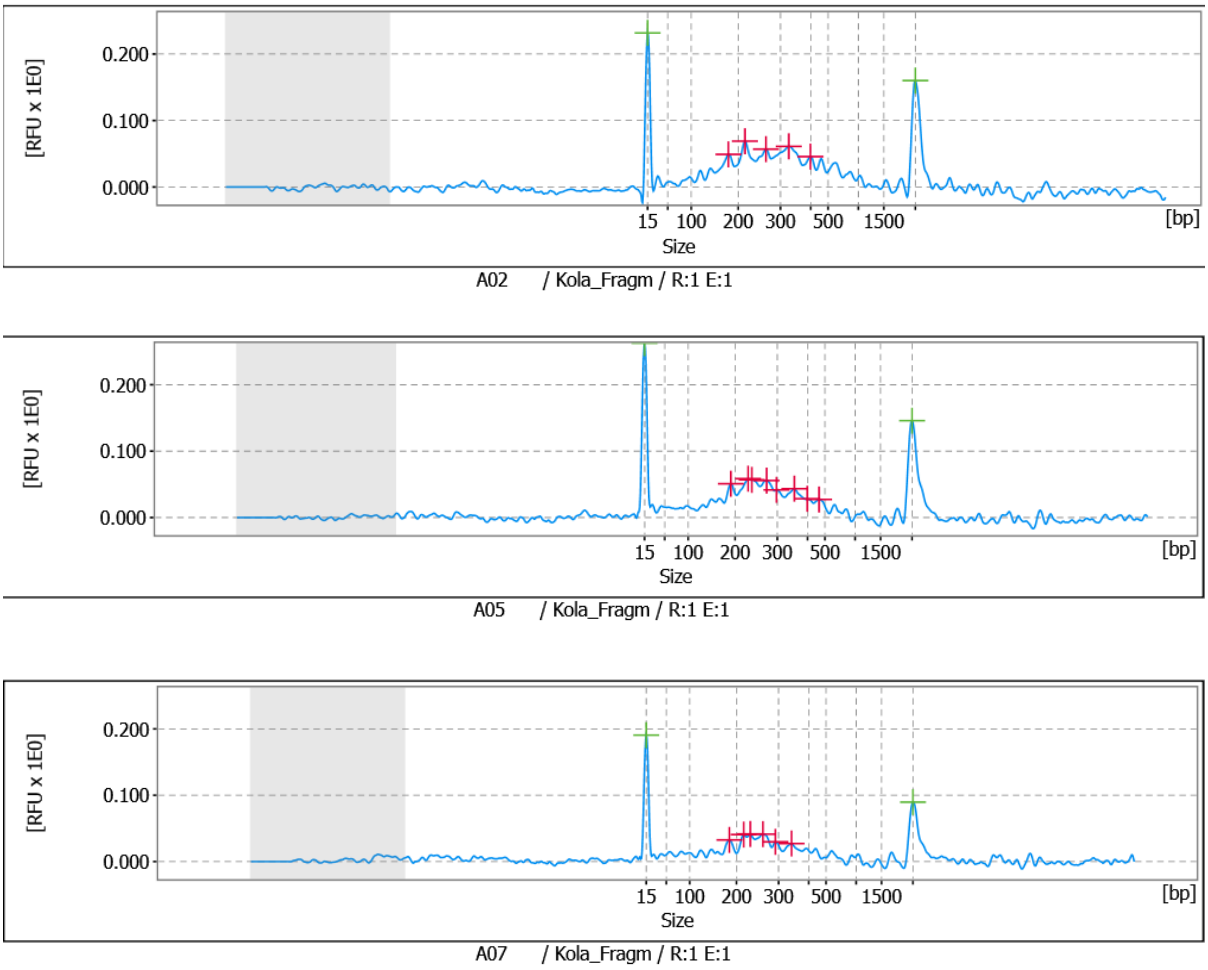

Plate 4.8: Final library size distribution of DNA samples

4.9 GENOME-SEQUENCING RUN SUMMARY

High quality sequence data obtained from 100 *P. falciparum* clinical isolates. Genome-wide short-read sequences were generated yielding a median read length of 182bp. Four hundred and eighty eight million bases were sequenced. Alignment to *P. falciparum* 3D7 reference genome indicated a 5.7x average coverage depth (Table 4.7). Variant calls were made at 56, 784 polymorphic sites and 13, 784 SNPs were identified after quality filtering. There was read coverage in the genic regions > 65% than in the intergenic regions. Five thousand one hundred and twenty two genes were analysed after excluding genes in the subtelomeric regions (var, rifin and stevor genes). Genes with at least 3 SNPs were considered informative for frequency-based analysis.

4.10 EVIDENCE OF POSITIVE DIRECTIONAL SELECTION

Standardised integrated haplotype score (iHS) was used to examine evidence of recent directional selection. Fourteen shared iHS regions that had at least 2 SNPs with a score > 2.5 were identified. These regions contained genes that were likely to have been under strong directional selection. Two of such genes were chloroquine resistance transporter (CRT) located in chromosome 7 and multidrug resistance 1 (MDR1) located in chromosome 5 genes. These genes also had high frequencies of nonsynonymous mutations (Figure 4.8). There was a weak signature of selection in dihydrofolate reductase (DHFR) in chromosome 4 and MDR5 genes in chromosome 13 with only 2 and 3 SNPs respectively identified within the iHS window. However, there was no evidence of recent directional selection in dihydropteroate synthase (DHPS) gene noticeable by the absent of chromosome 8 in the regions of the parasite genome with iHS value > 2.5 (Table 4.8). There was also a major selective sweep on chromosome 6

which had 32 SNPs within the shared iHS region. However, the origin and the specific loci targeted by the selective sweep were unknown.

4.11 IDENTIFYING SIGNATURES OF BALANCING SELECTION

Tajima's D values were mostly negative with a mean value of -0.86 (Appendix 7). One hundred and twelve genes (3.59%) had positive Tajima's D values (Figure 4.9). Six out of the twelve genes that had previously been shown to have positive indices also had positive values (Table 4.9).

Table 4.7: Alignment summary of the sequence reads

Length of non-gap regions covered by reads	13226305
Length of 24bp unique regions of the reference	264683
Reference nucleotide composition	**A**: 37.87%, **C**: 13.35%, **G**: 14.62%, T: 34.16%
Reads nucleotide composition	**A**: 36.07%, **C**: 15.71%, **G**: 17.71%, **T**: 30.51%
Average depth across all non-gap regions	0.691
Average depth across 24bp unique regions	0.514
Number of reference sequences	132994
Length of reference sequences excluding gaps	53277940
Length of gaps in the reference sequences	1299037
Average coverage depth of reference sequence	5.7x

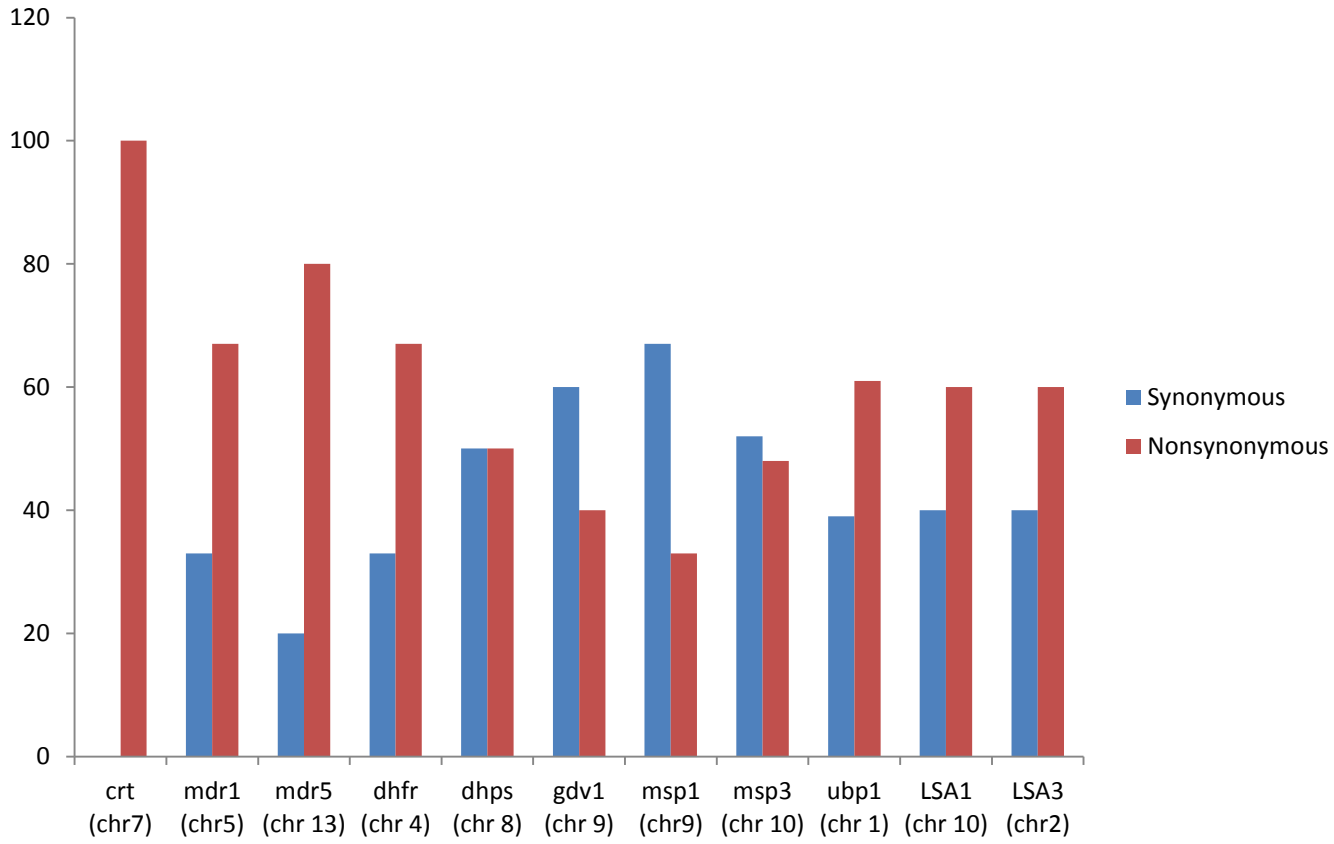

Figure 4.8: Distribution of synonymous and nonsynonymous SNPs among known gene targets of drug and immunity (CRT = Chloroquine resistant transporter; MDR = Multidrug resistant transporter; DHFR = dihydrofolate reductase; DHPS = dihydropteroate synthase; GDV = Gametocyte development; MSP = Merozoite surface protein; LSA = Liver stage antigen)

Table 4.8 Windows of directional selection for genomic regions with 2 or more SNPs and shared iHS score > 2.5

Chromosome	Window		Region (kb)	Number of SNPs	Id of genes within region
	Start	stop			
1	163	184	21	4	PF3D7_0103600 –PF3D7_0104200
2	302	515	213	1	PF3D7_0210900 –PF3D7_0213600
4	407	585	178	2	PF3D7_0409600 –PF3D7_0413300
5	822	1007	185	12	PF3D7_0519800 –PF3D7_0524200
6	**1039**	**1297**	**258**	**32**	**PF3D7_0625100 –PF3D7_0630600**
7	238	249	11	4	PF3D7_0704550 –PF3D7_0704900
9	1060	1220	160	15	PF3D7_0929320 –PF3D7_0930500
10	1210	1557	347	10	PF3D7_1029500 – PF3D7_103880
11	1190	1310	200	8	PF3D7_1132920 –PF3D7_1133600
12	1451	1880	429	6	PF3D7_1234900 –PF3D7_1248500
13	303	358	55	3	PF3D7_1306470 –PF33D7_130790
13	2410	2520	110	3	PF3D7_1306470 –PF3D7_1307900
14	1910	1932	22	2	PF37_1407000 – PF3D7_1447900

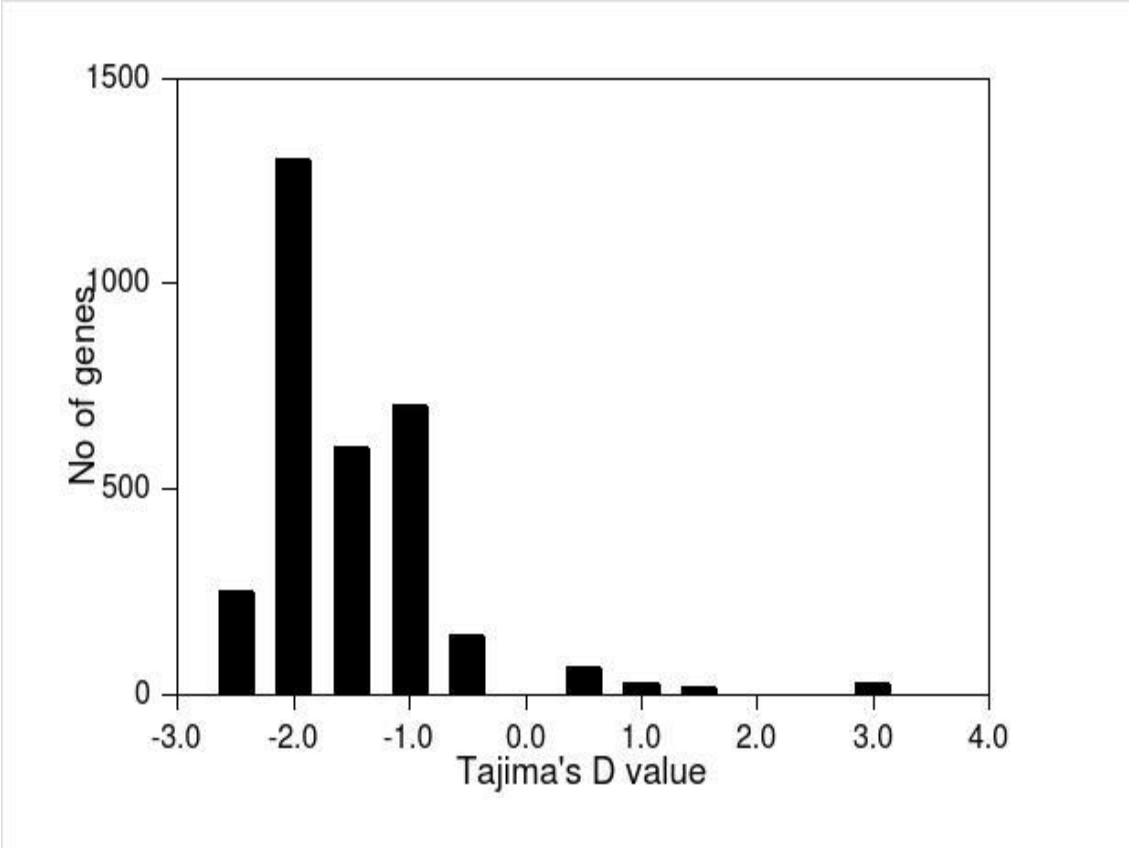

Figure 4.9: Frequency distribution histograms for individual gene values for Tajima's D

Table 4.9: Genes showing high values of Tajima's D

Gene ID	Chromosome no	Frequency of polymorphism π x 10^{-3}	Tajima's D
MSP(DBLMSP2)	10	27.9	2.58
CSP	3	20.6	1.38
EBA-175	7	17.2	1.29
MSP7	13	13.5	0.84
MSP3	10	23.8	0.73
SERA5	2	24.1	0.21

MSPDBL2=Merozoite surface protein duffy binding-like; CSP = circumsporozoite protein; EBA-175 = Erythrocyte-binding antigen; MSP = Merozoite Surface Protein; EBA = Erythrocyte-binding antigen; SERA = Serine repeat antigen

CHAPTER FIVE

5.0 DISCUSSION

Genetic investigation of malaria parasites provides vital information about the epidemiological patterns in a population and the possible existence of barriers that could limit gene flow between populations. In addition, varying selection caused by differences in host immunity and antimalarial drug pressure leads to evolutionary changes responsible for high level of genetic variations in the parasite. Hence effective control methods involve population-specific genomic studies to survey for genes under positive or equilibrium polymorphisms as a result of the influence of drug pressure or host immunity. This investigation was therefore timely as the country intensifies efforts on the control of malaria. The study employed microsatellites to determine the structure of *P. falciparum* populations and pooled sequencing (pool-seq) of the parasite genome using allele-frequency based neutrality test (Tajima's D) and integrated haplotype score (iHS) to identify genes under selection.

The occurrence of only *P. falciparum* in the parasite populations supports previous report that it is the predominant species responsible for malaria infection in Nigeria (WHO, 2014). More malaria cases were reported in females in conformity with the observations of Azfar *et al.* (2009). This may not be unconnected with the gender-related dynamics in treatment-seeking behaviour given that the study was hospital-based.

This is, nonetheless, not consistent with reports from Akanbi *et al*. (2010) who submitted that the production of more attractive chemicals for mosquitoes in males; higher levels of testosterone which is believed to suppress anti-plasmodial immune response in males (Krucken *et al*., 2005) and other genetic and hormonal factors (Zuk and McKean, 1996) may give females better

immunity against malaria infection. A large-scale community survey may be required to accurately understand gender norms and behaviour that impact the prevalence and severity of malaria infections. Also, the relatively high incidence of malaria reported in the rural community (Aramoko-Ekiti) suggests the need to scale-up malaria control efforts in the community and indeed the whole state which is often regarded as an orphan state due to limited interventions.

Besides the previously documented shortcomings of antigen-coding loci which include immune selection (Mwangi *et al*., 2006), this investigation has revealed the existence of few allelic variants of *P. falciparum* circulating within the population in southwestern Nigeria suggesting limited clonal diversity. Since molecular genotyping of *P. falciparum* antigenic markers are still being used to differentiate re-infection and recrudescence to adjust measures of treatment failure rates in antimalarial drug efficacy in Nigeria(NMEP, 2015), the implication of the existence of few allelic variants is that MSPs 1 and 2 may no longer be adequate for PCR correction of *in vivo* trials and outcomes because there is a strong probability of misclassifying some parasite infections, thereby overestimating treatment failure rates. This information is important even as research on the efficacy tracking of artemisinin-based combinations is being intensified (Gbotosho *et al*., 2011). Hence a modification of the standard MSP genotyping technique, the heteroduplex tracking assay, may be effective in tracking complex infections with high polyclonality. However, this assay depends on the use of radioactive tracers, making it difficult to employ in many field settings (Kwiek *et al*., 2007). Genotyping short sequence repeats (microsatellites) can also provide more accurate results given that they are more polymorphic and selectively neutral (Su *et al*., 1997). Designing a molecular barcode or signature for each given parasite to create a unique fingerprint (Daniels *et al*., 2008) that may not be compromised by high allelic frequencies can equally be utilised for malaria control in the country.

This study has revealed predominant distribution of MSP1 locus alleles belonging to the K1 family contrary to earlier reports from neighbouring states in Oyo (Amodu *et al*., 2005) and Ogun (Olasehinde *et al*., 2012). The finding does not also agree with reports from the North-Central region of the country (Oyedeji *et al*., 2013) where a higher proportion of parasite isolates possessing MSP2 alleles belonging to 3D7 family as against its FC27 counterpart, pointing to spatial dynamics in the genetic profile of *P. falciparum* populations in the country. This present study, however, corroborates the observations of Olasehinde *et al*. (2012) who described limited number of parasite genotypes in neighbouring southwestern communities. Given that the study involved cross-sectional sampling of clinical malaria parasites, there is a need to further understand the relationship between seasonality and diversity of the parasite populations taking into consideration the various ecological zones in the country.

The MOI values obtained in this study were low and they varied with populations ranging from 1.29 to 1.75. Previous investigation in southwestern Nigeria reported high MOI (Happi *et al.*, 2009) in contrast to this present report pointing to a trend in the reduction of malaria intensity since the multiplicity of an infection obtained by genotyping antigenic markers has an impact on the differentiation of recrudescence and re-infection as the sensitivity to detect multiple bands depends critically on the amount of DNA present in the sample (Ranford-Cartwright *et al*., 2002). It is also of particular interest since it is a proxy for intensity of infection, and allows predictions to be made about the potential for re-assortment and recombination in malaria populations owing to repeated sequential infection (Hill and Babiker, 1995) since high MOI favours recombination between genotypes and the breakdown of linkage disequilibrium.

Parasite clones within a human compete for resources and the removal of drug-sensitive clones following treatment allows surviving resistant clones to increase transmission in a process also known as competitive release (Wargo *et al*., 2007). Thus MOI is fundamental in the dynamics of the spread of drug resistance as it allows parasites with different resistant profiles to mate leading to non-random association between alleles encoding drug resistance. During the initial stages of spread of mutant alleles, high MOI values may entail a fast spread of resistant profiles but this process is reversed when resistant parasites approach fixation.

Hence the existence of low MOI may indicate a rapid spread of highly fit drug resistance alleles that have become fixed in the population. Earlier comparisons between MOI and age showed a greater MOI with age reflecting previous exposure to infection (Branch *et al*., 2001; Takala *et al*., 2007). Nonetheless, this study indicated that relatively high MOI values were reported among children 2-4 years old in accordance with a more recent finding by Agyeman-Budu *et al*. (2013).

The findings from this study indicated a low level of *P. falciparum* diversity with the allelic diversity values (0.65 -0.79) in the three parasite populations analysed tending towards the intermediate range. The heterozygosity values were lower than previously reported in other endemic countries with high levels of malaria transmission (Anderson *et al*., 2000; Conway *et al*., 2001; Joy *et al*., 2003). This suggests a possible decrease in *P. falciparum* transmission intensity. However, there is a need for persistent year-round monitoring for an adequate evaluation of success in control efforts. As intensity declines, transmission becomes increasingly focal leading to marked heterogeneity in malaria prevalence. These foci of transmission may not

only be due to environmental, climatic and ecological peculiarities (Getting *et al*., 2011), but also to the genetic uniqueness of the parasites creating hotspots consisting of a household or group of households maintaining higher transmission of malaria and a consistent reservoir of parasites throughout the year. Hence it important to characterize the genetic identities of the hotspots and hotpops of transmission in order to reduce local reservoirs of infection given that the scale-up of interventions may have reduced the disease burden in many endemic regions.

Population variance was determined with the analogue of Wright's Fst, AMOVA, and it was flexible enough to accommodate different types of assumptions about the evolution of microsatellites (Baloux and Lugon-Moulin, 2002). The analysis of molecular variance (AMOVA) value obtained for the populations sampled was low (0.017) indicating that almost all the genetic variations among parasites (98.3%) were contained within populations. Balloux and Lugo-Moulin (2002) put forward that population differentiation values of 0 – 0.05 suggest low genetic differentiation (GD) among populations; values between 0.05 - 0.15 could indicate moderate differentiation while higher values imply population partitioning into sub-groups. The low level of genetic differentiation observed in this present study is in agreement with reports from more widely separated but similarly endemic countries in West Africa (Anderson *et al*., 2000; Conway *et al*., 2001). However, they vary from values reported in parasite populations from less endemic Asian (Machado *et al*., 2004; Iwagami *et al*., 2009) and South American (Schultz *et al*., 2010) countries with similar geographical distances.

The low among population variance and the existence of an inverse relationship in the genetic and geographic distances between Badagry and Aramoko (AMK), separated by physical distance

of approximately 268km, may imply a relatively free gene flow of the malaria parasites across southwestern Nigeria. Population structure by site of sampling was also not evident by principal coordinate analysis though there was an insignificant sub-grouping distinguishable at the 2nd versus 3rd principal coordinates owing probably to experimental noise.

The absence of significant variance or population sub-structuring in the study areas eliminates the possibility of producing false-positive results for signals of selection in the parasite genome owing to population differentiation. Lack of sub-structuring also suggests that gene flow precludes local natural selection and genetic drift. This is expected as vector species distribution in southwestern Nigeria is also largely homogenous for *Anopheles gambiae* s.s negating any possibility of local selection by the vector species (Awolola *et al*., 2002; Onyabe *et al*., 2003). Similarities in the parasite populations may also be dependent on host variables such as human genetics. However, a national survey may further consolidate this observation as comparative parasitological, clinical and immunological surveys have shown consistent inter-ethnic differences in malaria infections (Modiano *et al*., 1996; Maiga *et al*., 2013). A more detailed investigation of the host-parasite genome-wide association studies in the ethnically diverse country may identify human and parasites genes important for local adaptation and disease manifestations.

In addition, it is important to quantify the association between resistance alleles especially when more than one locus is involved in drug resistance. Non-random association of resistance alleles at different loci has been shown to be a critical factor influencing the rate and spread of resistance (Dye and Williams, 1997).

In keeping with previous reports in other high malaria transmission areas with high levels of genetic recombination (Anderson *et al*., 2000), there was lack of significant linkage disequilibrium (LD) between markers in the three populations investigated. Though linkage was not significant for the entire sampled populations, parasite isolates from AMK with lower mean MOI had higher linkage disequilibrium. However, this finding does not agree with the theory which showed an association between transmission intensity and LD (Conway *et al.,* 2001) demonstrating higher levels of outbreeding in regions with higher transmission and that parasites from regions with low prevalence or low levels of multiple infections have higher levels of linkage disequilibrium than those from regions with high prevalence or with high levels of multiple infections (Mobegi *et al*., 2012).

This present report hence opens a new line of genetic investigation even as Hudson (1994) showed that LD is not directly comparable between populations when effective population size (Ne) varies. A theoretical framework to allow the interpretation of microsatellite-derived LD values in terms of levels of recombination will be useful. Although the linkage disequilibrium in the populations remains low compared to South-East Asia (Anthony *et al*., 2005; Iwagami *et al*., 2009), it will also be necessary to continue studies in this population and other isolated populations in Nigeria to detect any new patterns that may predict adaptation to drugs and host immunity.

Regions of the parasite genome under recent positive selection were identified, as they represented signatures of adaptation to drug pressure. Positive directional selection acting on a

beneficial trait gives rise to characteristic regions of low genetic diversity surrounding the causal genetic variant due to the preservation of linkage disequilibrium during meiosis (Su *et al*., 1999). Similar to earlier reports in other malaria endemic settings in West Africa (Chang *et al*., 2012; Mobegi *et al*., 2014), the existence of haplotype scores at high frequencies was demonstrated, occurring around chloroquine resistant transporter (CRT) and multidrug resistance (MDR1) genes.

The results of this study indicated that the population of resistant parasites did not reduce significantly years after drug use had been discouraged thus supporting the observation of high distribution of CQ-resistant parasites in western Kenya after a period of drug withdrawal (Zhong *et al.,* 2008). This finding is against the theoretical background of parasite population genetics which assumes that mutations encoding drug resistance pay a fitness penalty and are deleterious in the absence of the drug (Walliker, 2005). Hence the removal of drug pressure exposes resistant parasites to increased competition and should naturally lead to a decline in the frequency of resistance conferring selections as reported in Malawi where after just a decade of non-use, chloroquine (CQ) cleared 100% of asymptomatic *P. falciparum* infections (Kublin *et al*., 2003). Similarly, the prevalence of CQ-resistant parasites in coastal Tanzania decreased after only two and half years of CQ withdrawal (Temu *et al*., 2006).

The observations of strong selection in the CQ resistance genes in this present investigation may not be unrelated to long periods of chloroquine use leading to an absence of refugia (i.e. chloroquine-susceptible parasites) before the drug was finally replaced with ACT (FMoH, 2005). In populations where malaria is endemic, transmission is high, and acquired immunity is extensive, asymptomatic adults, who rarely become ill and who generally do not receive therapy,

may provide such refugia for susceptible parasites to persist in the population. In addition, malaria intervention efforts in Nigeria might have resulted in lower transmission and less acquired immunity, shrinking refugia and allowing resistant alleles to become fully fixed in the population, as has been observed in both South America (Cortese *et al*., 2002) and Southeast Asia (Nash *et al*., 2005).

Another plausible explanation for the enduring directional selection around CRT and MDR1 genes is the continued use of amodiaquine (AQ), a component of the newly introduced artemisinin-based therapy. AQ belongs to the group of 4-aminoquinolines and it is structurally similar to CQ. Association studies (Sa *et al*., 2009), showing cross-resistance between CQ and AQ may suggest that the use of artesunate plus amodiaquine contributes to continued persistence of resistance genotypes. Strong directional selection in CRT gene essentially implies that the re-introduction of the drug for malaria treatment cannot be considered yet as drug resistant strains still predominate the population. This is contrary to previous observations which showed complete (Kublin *et al*., 2003) or partial (Temu *et al*., 2006) reversal of *P. falciparum* populations to chloroquine susceptibility after a period of drug withdrawal.

Furthermore, there was a weak signature of selection in dihydrofolate reductase (DHFR) in chromosome 4 and MDR5 genes in chromosome 13 with only 2 and 3 SNPs respectively identified within the iHS window. There was no evidence of recent directional selection in dihydropteroate synthase (DHPS) gene noticeable by the absent of chromosome 8 in the regions of the parasite genome with iHS value > 2.5. In essence, the finding has revealed that there was no strong evidence of selection on DHFR and DHPS, the two major genes targeted by

sulfadoxine-pyrimethamine (sulfadox). This provides genetic evidence that sulfadox is still considerably effective for malaria chemoprophylaxis in concordance with the reports of Omole and Onademuren (2010). Thus its prophylactic use in Intermittent Preventive Treatment (IPT) of pregnant women may still be effective as observed by Adebayo *et al*. (2011). However, the detection of weak selection on DHFR gene observed in this study may be a pointer to increasing selection forces in the drug sites and resistance may emerge with time.

The study reports the distribution of synonymous (d_S) and non-synonymous mutations (d_{NS}) among some of the known target genes of drug and immunity revealing the higher levels of d_{NS} than d_S in the entire population, supporting previous observations by Amambua-Ngwa *et al*. (2012). This implies that the mutations or changes occurring in such genes may alter the amino acid sequence of the coded protein leading to important biological changes (resistance development or immune selection) in the parasite. It can be inferred that the genes are non-functional and are subject to natural selection.

This investigation has identified a major selective sweep on chromosome 6 of the parasite genome which was earlier described by Amambua-Ngwa *et al*. (2009) who targeted 75000 SNPs in parasite isolates from clinical cases in Senegal and The Gambia. However, the origin of the selective sweep is yet unknown and the specific loci targeted remain undefined. Further studies are, therefore, required to identify polymorphic microsatellite loci around the chromosome 6 sweep.

This study identified six genes to be under strong balancing selection (MSPDBL2, CSP, EBA-175, MSP3, MSP7 and SERA5). These were some of the genes that had been proposed as vaccine candidates from population genomic analyses of *P. falciparum* isolates between regions transiting from moderate to high endemicity southwards towards the Atlantic Coast (Amambua-Ngwa *et al*., 2012; Mobegi *et al*., 2014).

The gene with the highest Tajima's D value overall was the MSPDBL2 that also had the highest value among the candidate genes studied in East Africa (Ochola *et al*., 2010). MSPDBL2 is a merozoite protein incorporating the duffy-binding-like domains. It is, however, not only a target of immunity as functional validation has associated overexpression of the genes with lowered parasite sensitivity to both halofantrine and structurally related antimalarials such as mefloquine and lumefantrine. A single nucleotide polymorphism in codon 591 (C591S) of the gene has been shown to be involved in drug insensitivity (Ochola-Oyer *et al*., 2015). The observation of equilibrium polymorphism in this gene may have important implications on the efficacy of malaria treatment with artemisinin-based combinations. However, it remains to be proven if the C591S polymorphism can be used as a surveillance marker of lumefantrine resistance in the field.

Positive Tajima's D value of circumsporozoite protein as demonstrated by this study has provided population-specific evidence on the deployability of RTS, S vaccine (Mosquirix^{TM}) in the country. Circumsporozoite protein, a secreted protein of the sporozoite stage of the malaria parasite, is the major component of the pre-erythrocytic RTS, S vaccine that has passed through the third phase of clinical trials (WHO, 2015). Although the vaccine trials did not adequately

take into consideration parasite populations in Nigeria, the genomic information obtained from this study revealed signatures of balancing selection in the CSP suggesting that the antigen is under equilibrium polymorphism and may be stable over a very long period of time. However, the finding does not explain away the need to carry out trials on the immunogenicity and tolerability of the vaccine especially in an endemic area as Nigeria.

Promises and prospects notwithstanding, signatures of balancing selection may be transient and the efficacy of selection may be low even if selection coefficients are high (Connallon and Clark, 2014). Balancing selection may also be widespread, but its most common form may be difficult to detect. Worse still, drifts may completely overwhelm the signal of balancing selection, making it effectively undetectable with population genetic methods (Pespeni *et al*., 2012). This clearly elucidates the difficulty in predicting the actual frequency of selection even as the host-parasite interactions change dynamically through time. Thus there is a need for independent and consistent temporal as well as spatial investigations of selections on *P. falciparum* populations.

Further studies are also required to understand the properties of balancing selection using model-based methods which may involve high throughput fitness assays (Burke *et al*., 2010) as well as genomic analyses of experimental evolution (Orozco-ter Wengel *et al*., 2012). The integration of the lines of evidence from the two approaches is expected to provide accurate frequency estimations and the prediction of fractions of balancing selection that are detectable over time.

Another caveat in applying genomic techniques is the possible occurrence of ancillary causes of selection besides immunity which may operate on the parasite at other stages of development. This may include rapidly increased frequency of antigenic alleles at a locus which occasionally

happens per chance and may be indistinguishable from selection events. As well as maintaining diversity in selected sites, balancing selection increases diversity in flanking neutral sites (Hudson and Kaplan, 1988; Charlesworth *et al*., 1997; Takahata and Satta, 1998). Alleles existing at a locus for a period of time can acquire individual sets of neutral mutations that are unique unless recombination events occur causing a linkage between the functionally different alleles. The region around these alleles where recombination events have occurred can thus differ at multiple non-selected sites so that polymorphism will be higher than in unlinked genome regions, over a distance depending on regional recombination frequency (Charlesworth, 2006).

In *P. falciparum*, patterns of polymorphism induced by balancing selection are localized in individual genes (Tetteh *et al*., 2009; Amambua-Ngwa *et al*., 2012). This phenomenon makes it difficult to distinguish between long-term balancing selection with recombination and temporary maintenance of alleles as exchanges between alleles become more frequent within localized genes. Also, heterozygote advantage operating at the diploid stage of the parasite in the midgut of the *Anopheles* mosquito or non-self recognition among heterologous asexual stages of the parasites can pose some difficulties in inferring selection in human populations (Amambua-Ngwa *et al*., 2012).

Given the challenges of population subdivision and history associated with MK, HKA and Tajima's D, established tests for selection require supplementary evidence. A shorter time-line between antigen discovery and clinical trials through adequate funding and scientific commitment is also important in designing a largely protective malaria vaccine.

CHAPTER SIX

6.0 SUMMARY OF FINDINGS

The following findings emanated from the genome-wide analysis and diversity of *P. falciparum* in Lagos and Ekiti States, southwestern Nigeria.

1. The prevalence of malaria was highest in Aramoko (38.5%), 26.1% in Badagry and lowest in Lekki (22.1%).

2. All the three families of MSP-1 (K1, MAD20 and RO33) and two of MSP-2 (FC27 and 3D7) were observed among the *P. falciparum* isolates from all the studied populations.

3. There were few allelic variants leading to low clonal diversity that was highest in AMK and lowest in BDG.

4. Highest and lowest mean multiplicity of infections (MOI) with varying clones of *P. falciparum* were recorded in AMK (1.76) and BDG (1.29) respectively. When the study populations were taken together, mean MOI was observed to be highest in individuals between 2-4 years (1.80) and lowest in individuals above 18 years.

5. Genetic diversity values were similar across all populations, with mean H_E values across all loci between 0.65 (for LEK) and 0.79 (for AMK). Although the mean number of genotypes detected per isolate was highest in AMK, there was no substantial difference in the mean number of genotypes in the three parasite populations.

6. Genetic differentiation between the studied populations was low and there was no observable relationship between genetic distance and the natural log of the geographical distance for each pair of parasite populations studied. There was also low-level of population sub-structuring of the parasite within the region.

7. Pooled DNA sequencing provided appreciable information on the genome-wide patterns of selection of *P. falciparum* previously achievable through individual sequencing, identifying regions of the parasite genome under the influence of drug and host immunity.

8. Chloroquine resistance transporter (CRT) and multidrug resistance 1 (MDR1) genes (the sites associated with resistance to chloroquine) are under strong directional selection.

9. There was a weak signature of selection in dihydrofolate reductase (DHFR) and no evidence of recent selection in dihydropteroate synthase (DHPS) gene. Both genes are associated with resistance to sulfadoxine-pyrimethamine.

10. Merozoite surface proteins (MSP3 and MSP7), serine repeat antigen (SERA-5), erythrocyte-binding antigen (EBA-175), circumsporozoite (CSP) and merozoite surface protein duffy binding-like (MSPDBL2) were the genes with the highest Tajima's D values.

6.1 CONTRIBUTIONS TO KNOWLEDGE

1. This investigation has identified MSP3, MSP7, SERA-5 EBA-175, CSP and MSPDBL2 as potential target antigens of host immunity which can serve as vaccine candidates.

2. The study shows that CRT and MDR1 genes associated with resistance of *P. falciparum* to CQ are under drug pressure; hence the re-introduction of CQ will be inappropriate.

3. There is no genetic evidence to show that DHFR and DHPS genes are under strong selection pressure therefore use of SP for chemoprophylaxis should not be discontinued despite resistance concerns.

4. The study has shown low diversity and multiplicity of *P. falciparum* infections pointing to reduction in transmission intensity which suggests that the present malaria interventions may be yielding results.

5. The study has revealed low-level of genetic differentiation of *P. falciparum* parasites in southwestern Nigeria demonstrating that uniform chemotherapeutic and vaccinological approaches can be adopted for the region.

6.2 CONCLUSION

This investigation has provided information regarding the genetic diversity of *P. falciparum* populations in two southwestern states of Nigeria revealing no detectable structure between the populations and implying that uniform chemotherapeutic and vaccinological malaria intervention can be adopted for the region. The study also demonstrated the use of pooled DNA sequencing to understand genome-wide patterns of selection on *P. falciparum* previously achievable through individual sequencing thus identifying regions of the parasite genome under the impact of selection from antimalarial drug and predicting antigens that can serve as potential vaccine candidates.

6.3 RECOMMENDATIONS

In view of the molecular typing and genomic analysis of *P. falciparum* in Lagos and Ekiti States, southwestern Nigeria, the following recommendations were made to assist in malaria drug efficacy testing and vaccine construct:

1. The relatively high prevalence of malaria reported in the rural community (Aramoko-Ekiti) suggests the need to scale-up malaria control efforts in the community and indeed the whole state.

2. This study has revealed the existence of few allelic variants of *P. falciparum* in Lagos and Ekiti States emphasizing the need to review the existing methodology used in categorizing recrudescent and new infections during follow-up of antimalarial drug resistance tracking. This will prevent the over-estimation of treatment failure rates of antimalarial drugs.

3. The absence of a detectable population structure of *P. falciparum* in Lagos and Ekiti States implies that a fairly uniform malaria control strategy targeted at the parasites can be adopted in this highly endemic region.

4. The existence of strong selective forces in the two major genes associated with resistance of *P. falciparum* to chloroquine despite the official withdrawal of the drug for malaria treatment suggests continued drug pressure. Effective measures should be put in place to discourage the manufacture, importation, distribution, sale and use of chloroquine as an antimalarial. Continued surveillance will be useful to check whether resistance has decreased significantly.

5. Immunological analyses of the allelic protein products of the target antigens that have been identified as potential malaria vaccine candidates should now be prioritized.

REFERENCES

Abdel-Latif, M. S., Khattab, A., Lindenthal, C., Kremsner, P.G. and Klinkert, M.Q (2002). Recognition of variant Rifin antigens by human antibodies induced during natural *Plasmodium falciparum* infections. *Infectious Immunology*, **70** (12):7013-21.

Abdel-Latif, M. S., Cabrera, G., Kohler, C., Kremsner, P.G. and Luty, A.J. (2004). Antibodies to rifin: a component of naturally acquired responses to *Plasmodium falciparum* variant surface antigens on infected erythrocytes. *American Journal of Tropical Medicine and Hygiene*, **71** (2):179-86.

Abu-Raddad, L.J. and Ferguson, N.M. (2004). Improving epidemic malaria planning, preparedness and response in Southern Africa. *Malaria Journal*, **3**: 37.

Adams, J.H., Sim, B.K., Dolan, S.A., Fang, X., Kaslow, D.C. and Miller, L.H (1992). A family of erythrocyte binding proteins of malaria parasites. *Proceedings of the National Academy of Science*, USA, **89**: 7085-7089.

Adebayo, N.S., Oladele, H.A., Olalekan, A.W and Akindele, R.A (2011). The efficacy of sulphadoxine-pyrimethamine as a malaria prophylactic drug among pregnant women in Oshogbo, Southwestern Nigeria. *International Journal of Research and Reviews in Applied Sciences*, **3**: 42-52.

Adeoye, G.O and Nga, I.C (2007). Comparison of quantitative buffy coat technique (QBC) with Giemsa-stained thick film (GTF) for diagnosis of malaria. *Parasitology International*, **56** (4): 308-312.

Agyeman-Budu, A., Brown, C., Adjei, G., Adams, M., Dosco, D., Dery, D., Wilson, M., Asante,

K.P., Greenwood, B and Owusu-Agyei, S (2013). Trends in multiplicity of *Plasmodium falciparum* infections among asymptomatic residents in the middle belt of Ghana. *Malaria Journal*, **12**: 22.

Ahlborg, N., Berzins, K. and Perlmann, P. (1991). Definition of the epitope recognized by the *Plasmodium falciparum*-reactive human monoclonal antibody 33G2. *Molecular and Biochemical Parasitology*, **46**: 89-96.

Akanbi, O.M., Odaibo, A.B. and Ademowo, O.G. (2010). Effect of antimalarial drugs and malaria infection on oxidative stress in pregnant women. *African Journal of Reproductive Health*, **14**: 209-212.

Akey, J.M (2009). Constructing genomic maps of positive selection in humans: where do we go from here? *Genome Reserach*, **19**: 711-722.

Amambua-Ngwa, A. Tetteh, K.K., Manske, M., Gomez-Escobar, N., Stewart, L.B., Deerharke, M.E., Cheeseman, I.H., Newbold, C.I., Holder, A.A., Kneupfer, E., Janha, O., Jallow, M., Campino, S., Maclnnis, B., Kwiatwoski, D.P and Conway, D.J. (2012). Population genomic scan for candidate signatures of balancing selection to guide antigen characterization in malaria parasites. *PLOS Genetics*, **8**: e1002992

Amarantunga, C., Streng, S., Suon, E.S., Phelps, K., Stepniewska, P. and Zhou, C (2012). Artemisinin-resistant *Plasmodium falciparum* in Pursat province, Wetern Cambodia; a parasite clearance rate study. *The Lancet Infectious Diseases*, **12**: 851-858.

Amodu, O.K., Gbadegesin, R.A., Ralph, S.A., Adeyemo, A.A., Brenchley, P.E., Ayoola, O.O., Orimadegun, A.E., Akinsola, A.K., Olumese, P.E and Omotade, O.O. (2005) *Plasmodium falciparum* malaria in south-west Nigerian children: is the polymorphism of

ICAM-1 and E-selectin genes contributing to the clinical severity of malaria? *Acta Tropica*, **95** (3): 248-255.

Anders, R. F. (1986). Multiple cross-reactivities amongst antigens of *Plasmodium falciparum* impair the development of protective immunity against malaria with special reference to oxidant stress, *Parasite Immunology*, **8**: 529-539.

Anderson, T.J., Su, X.Z., Bockarie, M., Lagog, M. and Day, K.P. (1999). Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger-prick blood samples. *Parasitology*, **119**: 113-125.

Anderson, T.J., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J., Velez, I.D., Brockman, A.H., Nosten, F., Ferreira, M.U. and Day, K.P. (2000). Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Molecular Biology and Evolution*, **17**: 1467–1482.

Anderson, T.J.C., Polley, S.D., Vogler, A.P and Conway, D.J. (2007). Evidence of non-neutral polymorphism in *Plasmodium falciparum* gamete surface protein genes Pfs47 and Pfs48/45. *Molecular and Biochemical Parasitology*, **156**: 177-123.

Anthony, T., Conway, D., Cox-Singh, J., Matusop, A., Ratnam, S., Shamsul, S. and Singh B (2005). Fragmented population structure of *Plasmodium falciparum* in a region of declining endemicity. *Journal of Infectious Diseases*, **191**:1558–1564.

Arakawa, T., Komesu, A. and Otsuki, H (2005) Nasal immunization with a malaria transmission-

blocking vaccine candidate, Pfs25, induces complete protective immunity in mice against field isolates of *Plasmodium falciparum*. *Infectious Immunology*, **73**: 7375–80.

Arevalo-Herrera, M., Solarte, Y. and Yasnot, M.F (2005) Induction of transmission-blocking immunity in Aotus monkeys by vaccination with a *Plasmodium vivax* clinical grade PVS25 recombinant protein. *American Journal of Tropical Medicine and Hygiene*, **73**: 32–7.

Ariey, F., Witkowski, B., Amarantunga, C., Beghain, J., Langlois, A.C., Khim, N., Kim, S., Duru, V., Bouchier, C., Ma, L., Lim, P. and Menard, D. (2014). A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature*, **505** (7481): 50-5.

Audran, R., Cachat, M., Lurati, F., Soe, S., Leroy, O., Corradin, G., Druilhe, P. and Spertini, F. (2005). *Infectious Immunology*, **73** (12): 8017-26.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Miller, J.A., Nayak, V., Pennington, C., Pinney, D.F., Roos, D.S., Ross, C., Stoeckert, C.J., Jr., Treatman, C and Wang, H. (2009). PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research*, **37**: D539-543.

Awolola, T.S., Brooke, B.D., Hunt, R.H. and Coetzee, M (2002). Resistance of the malaria vector *Anopheles gambiae* s.s to pyrethroid insecticide in southwestern Nigeria. *Annals of Tropical Medicine and Parasitology*, **96**:849–852.

Azfar, F., Qayyum, A., HaleenM, A. and Ghaffar, A. (2009). Haematological abnormalities in malaria. *Biomedica*, **25**: 52-55.

Ballou, W.R. and Cahill, C.P. (2007). Two decades of commitment to malaria vaccine

development: GlaxoSmithKline Biologicals. *American Journal of Tropical Medicine and Hygiene*, **77**: 289-295.

Balloux F and Lugon-Moulin N (2002). The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**: 155–165.

Barber, B., William, T., Jelip, J., Rahman, H., Ibrahim, Y., Menon, J., Grigg, M., Yeo, T. and Anstey, N. (2011). Increasing incidence of *Plasmodium knowlesi* malaria following control of *P. falciparum* malaria in Sabah, Malaysia. *Menzies School of Health Research*.

Battle, K.E., Gething, P.W., Elyazar, I.R., Moyes, C.L., Sinka, M.E., Howes, R.E., Guerra, C.A., Price, R.N., Baird, K.J. and Hay, S.I. (2012). The global public health significance of *Plasmodium vivax*: timing of recurrence and whole blood chloroquine levels. *American Journal of Tropical Medicine and Hygiene*, **56**: 621-626.

Berzins, K. and Anders, R.F. (1999). The malaria antigens. In Wahlgren, M. and Perlmann, P. (ed.), Malaria-molecular and clinical aspects. *Harwood Academic Publishers*, The Netherlands.

Blackman, M.J., Heidrich, H.G., Donachie, S., McBride, J.S. and Holder, A.A. (1990). A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies. *Journal of Experimental Medicine*, **172**: 379-382.

Blackman, M.J., Scott-Finnigan, T.J., Shai, S. and Holder, A.A. (1994). Antibodies inhibit the protease-mediated processing of a malaria merozoite surface protein. *Journal of Experimental Medicine*, **180**: 389-393.

Bonelo, A.D., Valmori, F., Triponez, J.M., Tiercy, G., Mentha, J., Oberholzer, P., Champagne, J.F., Romero, F., Esposito, I., Nebie, C., Barbey, P., Romero, S., Herrera, G., Corradin, G. and Lopez, J.A. (2000). Generation and characterization of malaria-specific human $CD8^+$ lymphocyte clones: effect of natural polymorphism on T cell recognition and endogenous cognate antigen presentation by liver cells. *European Journal Immunology*, **30**: 3079-3088.

Borst, P. (2003). Mechanisms of Antigenic Variation: An Overview. in Antigenic Variation *Academic Press, Elsevier*:1-15.

Bowen, D.G., McCaughan, G.W. and Bertolino, P. (2005). Intrahepatic immunity: a tale of two sites? *Trends in Immunology*, **26**: 512-517.

Bowman, S., Lawson, D. Basham, D. Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B. and Kyes, S. (1999). The complete nucleotide sequence of chromosome 3 of P*lasmodium falciparum*. *Nature,* **400**: (6744):532-8.

Bozdech, Z., Llinas, M., Pulliam, B.L., Wong, E.D., Zhu, J. and DeRisi, J.L. (2003). The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum. PLoS Biology*, DOI: 10.1371

Branch, O.H., Takala, S. and Kariuki, S (2001). *Plasmodium falciparum* genotypes, low complexity of infection and resistance to subsequent malaria in participants in the Asembo Bay Cohort Project. *Infection and Immunity*, **69**: 7783-7792.

Branquinho, M.S., Lagos, C.B., Rocha, R.M., Natal, D., Barata, J.M., Cochrane, A.H., Nardin,

E., Nussenzweig, R.S. and Kloetzel, J.K. (1993). Anophelines in the state of Acre, Brazil infected with *Plasmodium falciparum*, *P. vivax*, the variant *P. vivax* VK247 and *P. malariae*. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, **87**: 391-394.

Brown, K. N. (1973). Antibody induced variation in malaria parasites. *Nature*, **242** (5392):49-50.

Bruce, M.C., Donnelly, C.A., Aplers, M.P., Galinski, M.R., Barnwell, D., Walliker, D. and Day, K.P (2006). Cross-species interactions between malaria parasites in humans. *Science*, **287**: 845-848.

Bull, P.C., Lowe, B.S., Kortok, M., Molyneux, C.S., Newbold, C.I. and Marsh, K. (1998). Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nature Medicine*, **4**: 358-360.

Burke, M.K., Dunham, J.P., Shahrestan, P., Thomton, K.R., Rose, M.R. and Long, A.D. (2010). Genome-wide analysis of a long term evolution experiment with Drosophila. *Nature*, **467**: 7587-7590.

Bustamante, L.Y., Bartholdson, S.J., Crosnier, C., Campos, M.G., Wanaguru, M. and Nguon, C. (2013). A full-length recombinant *Plasmodium falciparum* PfRH5 protein induces inhibitory antibodies that are effective across common PfRH5 genetic variants. *Vaccine*, **31**: 373-379.

Calderwood, M.S., Gannoun-Zaki, L., Wellems, T.E. and Deitsch. K.W., (2003). *Plasmodium falciparum* var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *Journal of Biology and Chemistry*, **278** (36): 34125-32.

Carter, R., Mendis, K.N., Miller, L.H., Molineaux, L and Saul, A (2000) Malaria transmission-blocking vaccines–how can their development be supported? *Nature Medicine*, **6**: 241–4.

Carvalho, L.J., Oliveira, S.G., Theisen, M., Alves, F.A., Andrade, M.C., Zanini, G.M., Brigido, M.C., Oeuvray, C., Povoa, M.M. and Muniz, J.A. (2004). Immunization of *Saimiri scuireus* monkeys with *Plasmodium falciparum* merozoite surface protein-3 and glutamate-rich protein suggests that protection is related to antibody levels. *Scandinavian Journal of Immunology*, **59**: 363-372.

Cavanagh, D.R., Elhassan, I.M., Roper, C., Robinson, V.J., Giha, H., Holder, A.A., Hviid, L., Theander, T.G., Arnot, D.E. and McBride, J.S. (1998). A longitudinal study of type-specific antibody responses to *Plasmodium falciparum* merozoite surface protein-1 in an area of unstable malaria in Sudan. *Journal of Immunology*, **161**: 347-359.

Chakrabarti, D., Reddy, G.R., Dame, J.B., Almira, E.C., Laipis, P.J., Ferl, R.J., Yang, T.P., Rowe, T.C. and Schuster, S.M. (1994). Analysis of expressed sequence tags from *Plasmodium falciparum. Molecular and Biochemical Parasitology*, **66**: 97-104.

Chang, H.H., Park, D.J., Galinsky, K.J., Schaffner, S.F., Ndiaye, D., Ndir, O., Mboup, S., Wiegand, R.C., Volkman, S.K., Sabeti, P.C., Wirth, D.F., Neafsey, D.E. and Hartl, D.L. (2012). Genomic sequencing of *Plasmodium falciparum* malaria parasites from Senegal reveals the demographic history of the population. *Molecular Biology and Evolution*, **29**: 3427-3439.

Charlesworth, B.M., Morgan, M.T. and Charlesworth, D. (1997). The effects of deleterious mutations on neutral molecular variation. *Genetics*, **134**: 1289-1303.

Charleworth, D. (2006). Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics*, DOI: 10.1371

Chen, Q., Heddini, A., Barragan, A., Fernandez, V., Pearce, F. and Wahlgren. M. (2000). The semiconserved head structure of *Plasmodium falciparum* erythrocyte membrane protein 1 mediates binding to multiple independent host receptors. *Journal of Experimental Medicine*, **192** (1):1-9.

Chukwuocha, U.M., Nwakwuo, G.C. and Mmrole, I. (2013). Artemisinin-based combination therapy: Knowledge and perceptions of patent medicine dealers in Owerri metropolis, Imo State, Nigeria and implications for compliance with current malaria treatment protocol. *Journal of Community Health*, **38** (4): 759-765

Coil, D.A., Vandersmissen, L., Ginevra, C., Jarraud, S., Lammertyn, E. and Anne, J. (2008). Intragenic tandem repeat variation between *Legionella pneumophila* strains. *Biomed Central Microbiology*, **8**: 218.

Connallon, T. and Clark, A.G. (2014). Evolutionary inevitability of sexual antagonism. *Proceedings of the Royal Society of Biological Sciences*, **281**: 20132123.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, T.D., Andrews, T.D., Barnes, C., Campbell, P., Fitzgerald, T., Hu, M., Ihm, C.H., Kristiansson, K., Macarthur, D.G., Macdonald, D.G., Macdonald, J.R., Onyiah, I., Pang, A.W., Robson, S., Stirrups, K., Valsesia, A., Walter, K., Wei, J., Tyler-Smith, C., Carter, N.P., Lee, C., Scherer, S.W. and Hurles, M.E. (2010). Origins and functional impact of copy number variation in the human genome. *Nature*, **464**(7289): 704-712.

Conway, D.J., Cavanagh, D.R., Tanabe, K., Roper, C., Mikes, Z.S., Sakihama, N., Bojang, K.A.,

Oduola, A.M.J., Kremser, P.G., Arnot, D.E., Greenwood, B.M. and McBride, J.S. (2000). A principal target of human immunity to malaria identified by molecular population genetic and immunological analyses. *Nature Medicine*, **6**: 689-692.

Conway, D.J., Machado, R.L.D., Singh, B., Dessert, P., Mikes, Z.S., Povoa, M.M., Oduola A.M.J and Roper, C (2001). Extreme geographical fixation of variation in the *Plasmodium falciparum* gamete surface protein gene Pfs48/45 compared with microsatellite loci. *Molecular and Biochemical Parasitology*, **115**: 145–156.

Conway, D.J. and Baum, J. (2002). In the blood-the remarkable ancestry of *Plasmodium falciparum*. *Trends in Parasitology*, **18**: 351-355.

Conway, D.J. and Polley, S.D. (2002). Measuring immune selection. *Parasitology*, **125**: S3–S16.

Cooper, K.A., Adelekan, D.A., Esimai, A.O., Northrop-Clewes, C.A. and Thurnham, D.I. (2002). Lack of influence of red palm oil on severity of malaria infection in pre-school Nigerian children. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **96**(2): 216-223.

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

Daly, T.M. and Long, C. (1995). Humoral immune response plays a predominant role in controlling blood stage infection in rodent malaria. *Infectious Immunology*, **61**: 2462-2467.

Daniels, R., Volkman, S.K., Miner, D.A., Mahesh, N., Neafsey, D.E., Park, D.J., Rosen, D., Angelino, E., Sabeti, P.C., Wirth, D.F. and Wiegand, R.C. (2008). A general SNP-based molecular barcodes for *Plasmodium falciparum* identification and tracking. *Malaria Journal*, **7**: 223.

Da Silveira, L.A., Dorta, M.L., Kinura, E.A., Katzin, A.M., Kawamoto, F., Tanabe, K. and Ferreeira, M.U. (1999). Allelic diversity and antibody recognition of *Plasmodium falciparum* merozoite surface protein 1 during hypoendemic malaria transmission in the Brazillian amazon region. *Infectious Immunology*, **67**: 5906-5916.

Dieckmann, U., Metz, J.A.J., Sabelis, M.W. and Sigmund, K. (eds) (2002). Virulence Management: The Adaptive Dynamics of Pathogen–Host Interactions. Cambridge University Press, Cambridge.

Djimdé, A., Doumbo, O.K., Cortese, J.F., Kayentao, K., Doumbo, S. and Diourté, Y. (2001). A molecular marker of chloroquine-resistant falciparum malaria. *New England Journal of Medicine*, **344**: 257-263.

Dondorp, A.M., Nosten, P., Yi, D., Das, A., Phyo, J., Tarning, K.M. and Lwin, E. (2009). Artemisinin resistance in *Plasmodium falciparum* malaria. *The New England Journal of Medicine*, **361**: 455-467.

Duraisingh, M. T., Voss, T.S., Marty, A.J., Duffy, M.F., Good, R.T., Thompson, J.K., Freitas-Junior, L.H., Scherf, A., Crabb, B.S. and Cowman, A.F. (2005). Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell*, **121** (1): 13-24.

Dye, C. and Williams, B.G. (1997). Multigenic drug resistance among inbred malaria parasites, *Proceedings of the National Academy of Science USA*, 7: 277-281.

Ellis, R.D., Martin, L.B., Shaffer, D., Long, C.A. and Miura, K. (2010). Phase 1 trial of the *Plasmodium falciparum* blood stage vaccine MSP-1(42)-C1/Alhydrogel with and without CPG 7909 in malaria naïve adults. *PLoS ONE*, **5**: e8787.

El Sahly, H.M., Patel, S.M., Atmar, R.L., Lanford, T.A., Dube, T., Thompson, D., Sim, B.K., Long, C. and Keitel, W.A. (2010). Safety and immunogenicity of a recombinant nonglycosylated erythrocyte binding antigen 175 region II malaria vaccine in healthy adults living in an area where malaria is endemic. *Clinical Vaccine Immunology*, **17**(10): 1552-1559.

Escalante, A.A., Lal, A.A. and Ayala, F.J. (1998). Genetic polymorphism and natural selection in the malaria parasite *Plasmodium falciparum*. *Genetics*, **149**: 189–202.

Esen, M., Kremsner, P.G., Schleucher, R., Gassler, M., Imoukhuede, E.B., Imbault, N., Leroy, O., Jepsen, S., Knudsen, B.W., Schumm, M., Knobolch, J., Theisen, M. and Mordmuller, B. (2009). Safety and immunogenicity of GMZ2 – an MSP3-GLURP fusion protein malaria vaccine candidate. *Vaccine*, **27**(49): 6862-6888.

Excoffier, L., Smouse, P.E. and Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**: 479–491.

Federal Ministry of Health (2005). National antimalarial treatment guidelines. *FMoH*, 5-22.

Federal Ministry of Health (2013). National malaria control programme. Strategic plan 2009-2013. *FMoH,* 12-15.

Felger, I., Marshal, V.M., Reeder, J.C., Hunt, J.A., Mgone, C.S. and Beck, H.P. (1997). Sequence diversity and molecular evolution of the merozoite surface antigen 2 of *Plasmodium falciparum*. *Journal of Molecular Evolution*, **45**: 154-160.

Fernandez, V., Hommel, M., Chen, Q., Hagblom, P. and Wahlgren. M. (1999). Small, clonally variant antigens expressed on the surface of the *Plasmodium falciparum*-infected erythrocyte are encoded by the rif gene family and are the target of human immune responses. *Journal of Experimental Medicine*, **190**: (10):1393-1404.

Ferreira, M.U., Ribeiro, W.L., Tonon, A.P., Kawamoto, F. and Rich, S.M. (2003). Sequence diversity and evolution of the malaria vaccine candidate merozoite surface protein-1 (MSP-1) of *Plasmodium falciparum. Gene*, **304**: 65-75

Ferreira, M.U. and Hartl, D.L. (2007). *Plasmodium falciparum*: worldwide sequence diversity and evolution of the malaria vaccine candidate merozoite surface protein-2 (MSP-2). *Experimental Parasitology*, **115**: 32–40.

Ferretti, L., Ramos-Onsins, S.E. and Perez-Enciso, M. (2013). Population genomics from pool sequencing. *Molecular Ecology*, **22**: 5561-5576.

Fidock, D.A., Nomura, T., Talley, A.K. and Cooper, R.A (2000). Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Molecule and Cell*, **6**: 861–871.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R.,

Bult, C.J., Tomb, J.F., Dougherty, B.A. and Merrick, J.M. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenza* Rd. *Science,* **269**(5223):496-512.

Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., Witney, A.A., Wolters, D., Wu, Y., Gardner, M.J., Holder, A.A., Sinden, R.E., Yates, J.R. and Carucci, D.J. (2002). A proteomic view of the *Plasmodium falciparum* life cycle. *Nature*, **419**(6906): 520-526.

Franks, S., Baton, L., Tetteh, K., Tongren, E., Dewin, D., Akanmori, B.D., Koram, K.A., Ranford-Cartwright, L. and Riley, E.M. (2003). Genetic diversity and antigenic polymorphism in *Plasmodium falciparum*: Extensive serological cross-reactivity between allelic variants of merozoite surface protein 2. *Infection and Immunity*, **71**(6): 3485-3495.

Fumagalli, M., Pozzoli, U., Cagliani, R., Comi, G.P. and Riva, S. (2009). Parasites represent a major selective force for interleukin genes and shape the genetic predisposition to autoimmune conditions. *Journal of Experimental Medicine*, **206**(6): 1395-1408.

Gamain, B., Langsley, G., Fourmax, M.N., Touzel, J.P., Camus, D., Dive, D. and Slomianny, C. (1996). Molecular characterization of the glutathione peroxidase gene of the human malaria parasite *Plasmodium falciparum. Molecular and Biochemical Parasitology*, **78**: 237-248.

Gamble, C., Ekwaru, P.J., Garner, P. and ter Kuile, F.O. (2007). Insecticide-treated nets for the prevention of malaria in pregnancy: A systematic review of randomized controlled trials. *PLoS Medicine*, **4**(3): e107

Gao, S., Basu, S., Yang, G., Deb, A. and Hu, M. (2013). Oral bioavailability challenges of natural products used in cancer chemoprevention. *Progressive Chemotherapy*, **25**: 1553-1575.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Rutherford, K., Chan, M.S., Martin, D.M., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barrell, B. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum. Nature*, **419**: 498-511.

Gbotosho, G.O., Sowunmi, A. and Happi, C.T. (2011). Kinetics of *Plasmodium falciparum* sex ratios: Application to evaluation of the potential of antimalarial drugs to influence malaria transmission. *The Open Tropical Medicine Journal*, **4**: 33-38.

Genton, B., Betuela, I., Felger, I., Al-Yaman, F., Anders, R.F., Saul, A., Rare, L., Baisor, M., Lorry, G., Brwon, V., Pye, D., Irving, D.O., Smith, T.A., Beck, H.P. and Alpers, M.P. (2002). A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea. *Journal of Infectious Diseases*, **185**: 820-827.

Gesase, S., Gosling, R.D., Hashim, R., Ord, R., Naidoo, I., Madebe, R., Mosha, J.F., Joho, A., Mandlia, V., Mrema, H., Mapunda, E., Savael, Z., Lemnge, M. and Chanramohan, D. (2009). High resistance of *Plasmodium falciparum* to sulphadoxine/pyrimethamine in Northern Tanzania and the emergence of dhps resistance mutation at codon 581. *PLoS ONE*, DOI: 10.1371/journal.pone.0004569.

Gething, P. W., Elyazar, I.R.F., Moyes, C., Smith, D.L., Battle, K.E., Guerra, C.A., Patil, A.P.,

Tatem., A.J., Howes, R.E., Myers, M.F., George, D.B., Horby, P., Wertheim, H.F.L and Hay, S.I. (2012). A long neglected world malaria map: *Plasmodium vivax* endemicity in 2010. DOI:10.1371/journal.pntd.0001814

Gibson, H.L., Tucker, J.E., Kaslow, D.C., Krettli, A.U., Colline, W.E., Kiefer, M.C., Bathurst, I.C. and Barr, P.J. (1992). Structure and expression of the gene for Pv200, a major blood-stage surface antigen of *Plasmodium vivax. Molecular and Biochemical Parasitology*, **50**: 325-333.

Gomes, A.C., Paula, M.B., Duarte, A.M., Lima, M.A., Malafronte, R.S., Mucci, L.F., Gotlieb, S.L. and Natal, D. (2008). Epidemiological and ecological aspects related to malaria in the area of influence of the lake at Porto Primavera dam, in western Sao Paulo state, Brazil. *Review of the Institute of Medicine and Tropical Sao Paulo*, **50**: 287-295.

Gonzalez-Ceron, L., Rodriguez, M.H., Nettel, J.A., Villareal, C., Kain, K.C. and Hernandez, J.E. (1999). Differential susceptibility of *Anopheles albinamus* and *Anopheles pseudopunctipennis* to infections with coindigenous *Plasmodium vivax* variants VK210 and VK247 in southern Mexico. *Infectious Immunology*, **67**: 410-412.

Gourdon, D.M (1997). The genetic structure of *Escherichia coli* populations in feral house mice. *Microbiology*, **143**: 2039–2046.

Greenwood, B.M. and Armstrong, J.R.M (1991). Comparison of two simple methods for determining malaria parasite density. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **85**: 186-188.

Griffin, J.T., Hollingsworth, T.D., Okell, L.C., Churcher, T.S., White, M., Bousema, T.,

Drakeley, C.J., Ferguson, N.M., Basanez, M.G. and Ghani, A.C. (2010). Reducing *Plasmodium falciparum* malaria transmission in Africa: a model-based evaluation of intervention strategies. *PLoS Medicine*, **7**(8): e1000324.

Gupta, S., Snow, R.W., Donnelly, C.A., Marsh, K. and Newbold, C. (1999). Immunity to non-cerebral severe malaria is acquired after one or two infections. *Nature Medicine*, **5**: 340-343.

Gupta, S. and Galvani, A. (1999). The effects of host heterogeneity on pathogen population structure. *Philosophical Transactions of the Royal Society of London Biological Science*, **354**(1384): 711-719.

Hall, N., Karras, M., Raine, J.D., Carlton, J.M., Kooij, T.W., Berriman, M., Florens, L., Janssen, C.S., Pain, A., Christophides, G.K., James, K., Rutherford, K., Harris, B., Harris, D., Churcher, C., Quail, M.A., Ormond, D., Doggett, J., Trueman, H.E., Mendoza, J., Bidwell, S.L., Rajandream, M.A., Carucci, D.J., Yates, J.R. 3rd, Kafatos, F.C., Janse, C.J., Barrell, B., Turner, C.M., Waters, A.P. and Sinden, R.E. (2005). A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic and proteomic analyses. *Science*, **307**(5706): 82-86.

Han, G.D., Zhang, X.J., Zhang, H.H., Chen, X.X. and Huang, B.C. (1999). Use of PCR/DNA probes to identify circumsporozoite genotype of *Plasmodium vivax* in China. *Southeast Asian Journal of Tropical Medicine and Public Health*, **30**: 20-23.

Happi, C.T., Gbotosho, G.O., Folarin, O.A. and Sowunmi, A. (2009). Selection of *Plasmodium*

falciparum multidrug resistance gene 1 alleles in asexual stages and gametocytes by artemether-lumefantrine in Nigerian children with uncomplicated falciparum malaria. *Antimicrobial Agents of Chemotherapy*, **53**: 888–895.

Hartl, D.L., Volkman, S.K., Nielsen, K.M., Barry, A.E., Day, K.P. and Wirth, D.F. (2002). The paradoxical population genetics of *Plasmodium falciparum. Trends in Parasitology*, **18**: 266-272.

Haubold, B., Travisano, M., Rainey, P.B. and Hudson, R.R (1998). Detecting linkage disequilibrium in bacterial populations. *Genetics*, **150**: 1341–1348.

Hawley, W.A., Phillips-Howard, P.A., ter Kulie, F.O., Terlouw, D.J., Vulule, J.M., Ombok, M., Nahlen, B.L., Gimnig, J.E., Kariuki, S.K., Kolczak, M.K. and Hightower, A.W. (2003). Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *American Journal of Tropical Medicine and Hygiene*, **68**(4 Suppl), 121-127.

Helmby, H., Cavelier, L., Pettersson, U. and Wahlgren, M. (1993). Rosetting *Plasmodium falciparum*-infected erythrocytes express unique antigens on their surface. *Infectious Immunology*, **61**:284-288.

Hill, W.G. and Babiker, H.A. (1995). Estimation of numbers of malaria clones in blood samples. *Proceedings of the Royal Society of London Biological Science*, **262**: 249-257.

Hinterberg, K., Scherf, A., Gysin, J., Toyoshima, T., Aikawa, M., Mazie, J.C., da Silva,

L.P. and Mattei, D. (1994). *P. falciparum*: the Pf332 antigen is sectreted from the parasite by a brefeldin A-dependent pathway and is translocated to the erythrocyte membrane via the Maurer;s clefts. *Experimental Parasitology*, **38**: 19-23.

Hisaeda, H., Stowers, A.W and Tsuboi, T (2000) Antibodies to malaria vaccine candidates Pvs25 and Pvs28 completely block the ability of *Plasmodium vivax* to infect mosquitoes. *Infectious Immunology*, **68**: 6618–23.

Holder, A.A. and Riley, E.M. (1996). Human immune response to MSP-1. *Parasitology Today*, **12**: 173-174.

Holder, A.A. and Blackman, M.J. (1994). What is the function of MSP-1 on the malaria merozoite? *Parasitology Today*, **10**(5): 182-184.

Hommel, M., David, P.H. and Oligino. L.D. (1983). Surface alterations of erythrocytes in *Plasmodium falciparum* malaria. Antigenic variation, antigenic diversity and the role of the spleen. *Journal of Experimental Medicine*, **157**:1137-1148.

Hudson, R.R (1994). Analytical results concerning linkage disequilibrium in models with genetic transformation and recombination. *Journal of Evolutionary Biology*, **7**: 535-548.

Hudson, R.R. and Kaplan, N.L. (1998). Statistical properties of the number of recombination events in the history of a smple of DNA sequences. *Genetics*, **111**: 147-164.

Hviid, L. (2005). Naturally acquired immunity to *Plasmodium falciparum* malaria in Africa. *Acta Tropica*, **95** (3):270-5.

Ikadai, H., Tanaka, H., Shibahara, N., Matsuu, A., Uechi, M., Itoh, N., Oshiro, S., Kudo,

N., Igarashi, I. and Oyamada, T. (2004). Molecular evidence of infections with *Babesia gibsoni* parasites in Japan and evaluation of the diagnostic potential of a loop-mediated isothermal amplification method. *Journal of Clinical Microbiology*, **42**: 2465-2469.

Iwagami, M., Rivera, P., Villacorte, E., Escueta, A., Hatabu, T., Kawazu, S., Hayakawa, T., Tanabe, K. and Kano S. (2009). Genetic diversity and population structure of *Plasmodium falciparum* in the Philippines. *Malaria Journal*, **8**: 96

James, S. P., Nicol, V.D. and Shute. P.G. (1932). *Proceedings of the Royal Society of Medicine*, **25**:1153-1186.

Jeffares, D.C., Pain, A., Berry, A., Cox, A.V. and Stalker, J. (2007). Genome variation and evolution of the malaria parasite *Plasmodium falciparum*. *Nature Genetics*, **39**: 120-125.

Jensen, A. T., Magistrado, P., Sharp, S., Joergensen, L., Lavstsen, T., Chiucchiuini, A. Salanti, A. Vestergaard, L.S., Lusingu, J.P., Hermsen, R. Sauerwein, R., Christensen, J., Nielsen, M.A., Hviid, L., Sutherland, C., Staalsoe, T. and Theander. T.G. (2004). *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A var genes. *Journal of Experimental Medicine*, **199**(9): 1179-90.

Jepsen, M.P.G., Jogdand, P.S., Singh, S.K., Esen, M., Christainsen, M., Issifou, S., Hounkpatin, A.B., Ateba-Ngoa, U., Kremsner, P.G., Dziegiel, M.H., Olesen-Larsen, S., Jepsen, S., Mordmuller, B. and Theisen, M. (2013). The malaria vaccine candidate GMZ2 elicits functional antibodies in individuals from malaria endemic and non-endemic areas. *Journal of Infectious Diseases*, **208**: 479-488.

Jeslyn, W.P., Huat, T.C., Vernon, L., Irene, L.M., Sung, L.K. and Jarrod, L.P. (2011).

Molecular epidemiological investigation of *Plasmodium knowlesi* in humans and macaques in Singapore. *Vector-Borne and Zoonotic Diseases*, **11**: 131–135.

Jongwutiwes, S., Buppan, P., Kosuvin, R., Seethmachai, S., Pattanawong, U. and Sirichaisinthop, J. (2011). *Plasmodium knowlesi* malaria in humans and macaques, Thailand. *Emerging Infectious Diseases*, **17**: 1799–1806.

Jordan, P., Snyder, L.A. and Saunders, N.J. (2003). Diversity in coding tandem repeats in related *Neisseria* spp. *Biomed Central Microbiology*, **3**: 23.

Joy, D.A., Feng, X., Mu, J., Furuya, T., Chotivanich, K., Krettli, A.U., Ho, M., Wang, A., White, N.J., Suh, E., Beerli, P., Su, X.Z. (2003). Reemergence of chloroquine-sensitive *Plasmodium falciparum* malaria after cessation of chloroquine use in Malawi. *Journal of Infectious Diseases*, **187**: 1870-1875.

Kain, K.C., Brown, A.E., Webster, H.K., Wirtz, R.A., Keystone, J.S., Rodriguez, M.H., Kinahan, J., Rowland, M. and Lanar, D.E. (1992). Circumsporozoite genotyping of global isolates of *Plasmodium vivax* from dried blood specimens. *Journal of Clinical Microbiology*, **30**: 1863: 1866

Kain, K.C., Brown, A.E., Lanar, D.E., Ballou, W.R. and Webster, H.K. (1993). Response of *Plasmodium vivax* variants to chloroquine as determinated by microscopy and quantitative polymerase chain reaction. *American Journal of Tropical Medicine and Hygiene*, **46**: 473-479.

Kaviratne, M., Khan, S.M., Jarra, W. and Preiser, P.R. (2002). Small variant STEVOR antigen is

uniquely located within Maurer's clefts in *Plasmodium falciparum*-infected red blood cells. *Eukaryotic Cell,* **1**(6): 926-935.

Kilama, W. and Ntoumi, F. (2009). Malaria: a research agenda for elimination era. *Lancet*, **374**: 1480-1482.

Kraemer, S. M. and Smith. J.D. (2003). Evidence for the importance of genetic structuring to the structural and functional specialization of the *Plasmodium falciparum* var gene family. *Molecular Microbiology,* **50** (5):1527-38

Krafts, K., Hempelmann, E. and Oleksyn, B. (2011). The colour purple: from royalty to laboratory, with apologies to Malachowski. *Biotechnology and Histochemistry,* **86**(1): 7-35.

Krucken, J., Dkhil, M.A., Braun, J.V., Schroetel, R.M., El-Khadragy, M. and Carmeliet, P. (2005). Testosterone suppresses protective responses of the liver to blood-stage malaria. *Infectious Immunology*, **73**: 436-443.

Kublin, J.G., Cortese, J.F., Njunju, E.M., Mukadam, R.A. (2003). Re-emergence of chloroquine-sensitive *Plasmodium falciparum* malaria after cessation of chloroquine use in Malawi. *Journal of Infectious Diseases,* **187**: 1870-1875.

Kumar, A., Sharma, V.P., Sumodan, P.K. and Thavaselvam, D. (1998). Field trials of biolarvicide *Bacillus thuringiensis* var. *israelensis* strain 164 and the larvivorous fish *Aplocheilus blocki* against *Anopheles stephensi* for malaria control in Gia, India. *Journal of American Mosquito Control Association*, 457- 462.

Kwiek, J.J., Alker, A.P., Wenink, E.C., Chaponda, M., Kakilani, L.V. and Meshnick, S.R.

(2007). Estimating true antimalarial efficacy by heteroduplex tracking assay in patients with complex *Plasmodium falciparum* infections. *Antimicrobial Agents and Chemotherapy*, **51**: 521-527.

Kyaw, M.P., Nyunt, M.H., Chit, K., Aye, M.M., Aye, K.H., Lindegardh, N., Tarning, J., Imwong, M., Jacob, C.G., Rasmussen, C., Perin, J., Ringwald, P. and Nyunt, M.M. (2013). Reduced susceptibility of *Plasmodium falciparum* to artesunate in Southern Myanmar. *PLoS ONE*, DOI: 10.1371/journal.pone.0057689.

Kyes, S.A., Rowe, J.A., Kriek, N. and Newbold, C.I. (1999). Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with *Plasmodium falciparum. Proceedings of the National Academy of Science USA*, **96**: 9333-9338.

Laufer, M.K., Harrison, T.S., Dzinjalamala, F.K. and Stine, O.C (2010). Return of chloroquine-susceptible falciparum malaria in Malawi was a re-expansion of diverse susceptible parasites. *Journal of Infectious Diseases*, **202**: 801–808.

Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De la Vega, P., Holder, A.A., Batalov, S., Carucci, D.J. and Winzeler, E.A. (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*, **301**: 1503-1508.

Le Roch, K.G., Chung, D.W. and Ponts, N. (2012) Genomics and integrated systems biology in *Plasmodium falciparum*: A path to malaria control and eradication. *Parasite Immunology*, **34**: 50-60.

Leech, J.H., Barnwell, J.W., Miller, L.H., Howard, R.J. (1984). Identification of a strain-specific

malarial antigen exposed on the surface of *Plasmodium falciparum*-infected erythrocytes. *Journal of Experimental Medicine*, **159**: 1567-1575.

Levdansky, E., Romano, J., Shadkchan, Y., Sharon, H., Verstrepen, K.J., Fink, G.R. and Osherov, N. (2007). Coding tandem repeats generate diversity in *Aspergillus fumigatus* genes. *Eukaryotic Cell*, **6**: 1380-1391.

Li, G.Q., Gupo, X.B., Fu, L.C., Jian, H.X. and Wang, X.H. (1994). Clinical trials of artemisinin and its derivatives in the treatment of malaria in China. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **88**: S5-6.

Ling, I.T., Cooksley, S., Bates, P.A., Hempelmann, E and Wilson, R.J.M (1986). Antibodies to the glutamate dehydrogenase of *Plasmodium falciparum*. *Parasitology*, **92** (2): 313-324.

Lipsitch, M. and O'Hagan, J.J. (2007). Patterns of antigenic diversity and the mechanisms that maintain them. *Royal Society Interface*, **4:**16

Locher, C.P., Tam, L.Q., Chang, S.P., McBride, J.S., Siddiqui, W.A. (1996). *Plasmodium falciparum*: gp 195 tripeptide repeat-specific monoclonal antibody inhibits parasite growth in vitro. *Experimental Parasitology*, **84**: 74-83.

Luse, S.A. and Miller, H.A. (1971). *Plasmodium falciparum* malaria. Ultrastructure of parasitized erythrocytes in cardiac vessels. *American Journal of Tropical Medicine and Hygiene*, **20**: 655-660.

Machado, R.L.D., Póvoa, M.M., Calvosa, V.S.P., Ferreira, M.U., Rossit, A.R.B., dos Santos, E.J.M. and Conway, D.J (2004). Genetic structure of *Plasmodium falciparum* populations in the Brazilian Amazon region. *Journal of Infectious Diseases*, **190**:1547–1555.

Mahanty, S., Saul, A. and Miller, L.H. (2003). Progress in the development of recombinant and synthetic blood-stage malaria vaccines. *Journal of Experimental Biology*, **206**: 3781-3788.

Maiga, B., Dolo, A., Toure, O., Dara, V., Tapily, A., Campino, S., Sepulveda, N., Risley, P., Silva, N., Corrran, P., Rockett, K.A. and Kwiatkowski, D (2013). Human candidate polymorphisms in sympatric ethnic groups differing in malaria susceptibility in Mali. *PLoS ONE*, DOI: 10.1371

Mancilla, L.I., Levitus, G., Kirchgatter, K., Mertens, F., Herrera, S. and del Portillo, H.A. (1994). *Plasmodium vivax*: dimorphic DNA sequences from the MSP-1-1 gene code for regions that are immunogenic in natural infections. *Experimental Parasitology*, **79**: 148-158.

March, S., Ng, S., Velmurugan, S., Galstian, A., Shan, J., Logan, D.J., Carpenter, A.E., Thomas, D., Sim, B.K. and Mota, M.M. (2013). A microsale huma liver platform that supports the hepatic stages of *Plasmodium falciparum* and vivax. *Cell Host Microbe*, **14**: 104-115.

Marsh, K. (1998). Malaria disaster in Africa. *Lancet*, **352**(9132): 924.

Marti, M., Good, R.T., Rug, M., Knuepfer, E. and Cowman, A.F. (2004). Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science*, **306**: (5703):1930-3.

McColl, D.J., Silva, A., Foley, M., Kun, J.F., Favaloro, J.M., Thompson, J.K., Marshall, V.M., Coppel, R.L., Kemp, D.J. and Anders, R.F. (1994). Molecular variation in a novel polymorphic antigen associated with *Plasmodium falciparum* merozoites. *Molecular and Biochemical Parasitology*, **68**(1): 53-67.

McColl, D.J. and Anders, R.F. (1997). Conservation of structural motifs and antigenic diversity

in the *Plasmodium falciparum* merozoite surface protein-3 (MSP-3). *Molecular and Biochemical Parasitology*, **90**(1): 21-31.

McCutchan, T.F., Piper, R.C. and Makler, M.T. (2008). Use of malaria rapid diagnostic test to identify *Plasmodium knowlesi* infection. *Emerging Infectious Diseases*, **14**(11): 1750-1752.

Mckenzie, F.E., Ferreira, M.U., Baird, J.K., Snounou, G. and Bossert, W.H. (2001). Meiotic recombination, cross-reactivity, and persistence in *Plasmodium falciparum*. *Evolution*, **55**: 1299-1307.

Menendez, C., Bardaji, A., Sigauque, B., Sanz, S., Aponte, J.J., Mabunda, S. and Alonso, P.L. (2010). Malaria prevention with IPTp during pregnancy reduces neonatal mortality. *PLoS ONE*, **5**(2): e9438.

Mens, P.F., Schoone, G.J., Kager, P.A. and Schallig, H.D.F. (2006). Detection and identification of human *Plasmodium* species with real-time quantitative nucleic acid sequence-based amplification. *Malaria Journal*, **5**:80.

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

Mita, T., Venkatessan, M., Ohashi, J., Culleton, R., Takahashi, N., Tsukahara, T., Ndounga, M., Dysoley, L., Endo, H., Hombhanje, F., Ferreira, M.U., Plowe, C.V. and Tanabe, K. (2011). Limited geographic origin and global spread of sulfadoxine-resistant dhps alleles in *Plasmodium falciparum* populations. *Journal of Infectious Diseases*, **204**: 1980-1988.

Mobegi, V.A., Loua, K.M., Ahoudi, A.D., Satoguina, J., Nwakanma, D.C., Amambua-Ngwa, A.

and Conway, D.J. (2012). Population genetic structure of *Plasmodium falciparum* across a region of diverse endemicity in West Africa. *Malaria Journal*, **11**: 223.

Mobegi, V. A., Duffy, C. W., Amambua-Ngwa, A., Loua, K. M., Laman, E., Nwakanma, D., Macinnis, B., Aspeling-Jones, H., Murray, L., Clark, T. G., Kwiatkowski, D. P. and Conway, D. J. (2014). Genome-wide analysis of selection on the malaria parasite *Plasmodium falciparum* in West African populations of differing infection endemicity. *Molecular Biology and Evolution*, **31** (6): 1490-1499.

Modiano, D., Petrarca, V., Sirima, B.S., Nebie, I., Diallo, D., Esposito, F and Colluzzi, M. (1996). Different response to *Plasmodium falciparum* malaria in west African sympatric ethnic groups. *Proceedings of the National Academy of Science USA*, **93**: 13206-13211.

Mohapatra, P. K., Prakash, A., Bhattacharyya, D. R., Goswami, B. K., Ahmed, A., Sarmah, B. and Mahanta, J. (2008). Detection and molecular confirmation of a focus of *Plasmodium malariae* in Arunachal Pradesh, India. *The Indian Journal of Medical Research*, **128** (1): 52–56.

Moorthy, V.S., Good, M.F. and Hill, A.V. (2004). Malaria vaccine developments. *Lancet*, **363**: 150-156.

Mori, Y., Nagamine, K., Tomita, N. and Notomi, T. (2001). Detection of loop-mediated isothermal amplification by turbidity derived from magnesium pyrophosphate formation. *Biochemical and Biophysical Research Communications*, **289**: 150-154.

Mu, J., Awadalla, P., Duan, J., McGee, K.M., Keebler, J., Seydel, K., McVean, G.A., Su, X.Z.

(2007). Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. *Nature Genetics*, **39**: 126–130.

Mu, J., Myers, R.A., Jiang, H., Liu, S., Ricklefs, S., Waisberg, M., Chotivanich, K., Wilairata, P., Krudsood, S., White, N., Udomsangpetch, R., Cui, L., Ho, M., Ou, F., Li, H., Song, J., Li, G., Wang, X., Seila, S., Sokunthea, S., Socheat, D., Strudevant, D.E., Porcelia, S.F., Fairhurst, R.M., Wellems, T.E., Awadaila, P. and Su, X. (2010). *Plasmodium falciparum* genome-wide scans for positive selection, recombination hotspots and resistance to antimalarial drugs. *Nature Genetics*, **42** (3): 268-271.

Mueller, A. K., Labaied, M., Kappe, S.H. and Matuschewski, K. (2005). Genetically modified *Plasmodium* parasites as a protective experimental malaria vaccine *Natur*e, **433** (7022): 164-7.

Murray, C.K. and Bennett, J.W. (2009). Rapid diagnosis of malaria. *Interdisciplinary Perspective on Infectious Diseases*. DOI: 10. 1155/2009/415953.

Mwangi, T.W., Bethony, J. and Brooker, S. (2006). Malaria and helminth interactions in humans: an epidemiological viewpoint. *Annals of Tropical Medicine and Parasitology*, **100** (7): 551-570.

Nagamine, K., Watanabe, K., Ohtuska, K., Hase, T. and Notomi, T. (2001). Loop-mediated isothermal amplification reaction using a non-denatured template. *Clinical Chemistry*, **47**: 1742-1743.

Nardin, E., Zavala, F., Nussenzweig, V. and Nussenzweig, R.S. (1999). Pre-erythrocytic malaria vaccine: mechanisms of protective immunity and human vaccine trials. *Parassitologia*, **41**: (1-3): 397-402.

Nash, D., Nair, S., Mayxay, M., Newton, P.N., Guthmann, J.P., Nosten, F., Anderson, T.J.C (2005). Selection strength and hitchhiking around two antimalarial resistance genes. *Proceeding of Biological Science*, **340**: 984-987.

National Malaria Elimination Programme (2015). *Malaria Operational Plan 2015*, 13-42

National Malaria Indicator Survey (2010). *NMIS*. Abuja, Nigeria

National Population Commission (2006). *NPC.* Abuja, Nigeria

Need, J.T., Wirtz, R.A., Fernandez, I., Franke, E.D., Rodriguez, M.H. and Lanar, D.E. (1993). *Plasmodium vivax* VK247 and VK210 circumsporozoite proteins in *Anopheles* mosquitoes from Andoas, Peru. *Journal of Medical Entomology*, **30**: 597-600.

Ng, O.T., Ooi, E.E., Lee, C.C., Lee, P.J., Ng, L.C. and Pei, S.W. (2008). Naturally acquired human *Plasmodium knowlesi* infection, Singapore. *Emerging Infectious Diseases,* **14**: 814–816.

Nielsen, M. A., Staalsoe, T., Kurtzhals, J.A., Goka, B.Q., Dodoo, D., Alifrangis, M., Theander, T.G., Akanmori, B.D. and Hviid, L. (2002). *Plasmodium falciparum* variant surface antigen expression varies between isolates causing severe and nonsevere malaria and is modified by acquired immunity. *Journal of Immunology*, **168** (7):3444-50.

Noedl, H., Se, K., Schaecher, K. and Smith, B.L. (2008). Evidence of artemisinin-resistant malaria in western Cambodia. *The New England Journal of Medicine*, **359**: 2619-2620.

Nussenzweig, R.S., Vanderberg, J., Most, H. and Orton, C (1967). Protective immunity produced by the injection of x-irradiated sporozoites of *Plasmodium berghei*. *Nature,* **216**: 160–2.

Ocholla, L.I., Tetteh, K.K., Stewart, L.B., Riitho, V., March, K. and Conway, D.J. (2010). Allele frequency-based and polymorphism-versus-divergence indices of balancing selection in a new filtered set of polymorphic genes in *Plasmodium falciparum*. *Molecular Biology and Evolution*, **27**: 2344-2351.

Ockenhouse, C.F., Betageri, R., Springer, T. and Staunton, D. (1992). *Plasmodium falciparum*-infected erythrocytes bind ICAM-1 at a site distinct from LFA-1, Mac-1 and human rhinovirus. *Cell*, **68**: 63-69.

O'Dushlaine, C.T., Edwards, R.J., Park, S.D. and Shields, D.C. (2005). Tandem repaeat copy-number variation in protein-coding regions of human genes. *Genome Biology*, **6**: R69.

Oeauvray, C., Bouharoun-Tayoun, H., Grass-Masse, H., Lepers, J.P., Ralamboranto, L., Tartar, A. and Druilhe, P. (1994). A novel merozoite surface antigen of *Plasmodium falciparum* (MSP-3) identified by cellular-antibody cooperative mechanism antigenicity and biological activity of antibodies. *Memorias do Instituto Oswaldo Cruz*, **89** suppl 2: 77-80.

Oeauvray, C., Theisen, M., Rogier, C., Trape, J., Jepsen, S. and Druilhe, P. (2000). Cytophilic immunoglobulin responses to *Plasmodium falciparum* glutamate-rich protein are correlated with protection against clinical malaria in Dielmo, Senegal. *Infectious Immunology*, **68**(5): 2617-2620.

Offeddu, V., Thathy, V., Marsh, K. and Matuschewski, K. (2012). Naturally acquired immune responses against *Plasmodium falciparum* sporozoites and liver infection. *International Journal of Parasitology*, **42**: 535–548.

Ofori, M. F., Dodoo, D., Staalsoe, T., Kurtzhals, J.A., Koram, K., Theander, T.G., Akanmori,

B.D. and Hviid, L. (2002). Malaria-induced acquisition of antibodies to *Plasmodium falciparum* variant surface antigens. *Infectious Immunology*, **70** (6):2982-8.

Ogutu, B.R., Apollo, O.J., Denise, M., Okoth, W., Siangla, J., Dubovsky, F., Tucker, K., Waitumbi, J.N., Diggs, C., Wittes, J., Elissa M., Soisson, L.A. and Withers, M.R. (2009). Blood stage malaria vaccine eliciting high antigen-specific antibody confers no protection to young children in Western Kenya. *PLoS ONE*, DOI:10.1371/journal.pone.0004708.

Oh, S.S., Voigt, S., Fisher, D., Yi, S.J., LeRoy, P.J., Derick, L., Liu, S.C. and Chishti, A.H. (2000). *Plasmodium falciparum* erythrocyte membrane protein 1 is anchored to the actin-spectrin junction and knob-associated histidine-rich protein in the erythrocyte skeleton. *Molecular and Biochemical Parasitology*, **108**: 237-247.

Okie S. (2005) Betting on a malaria vaccine. *New England Journal of Medicine*, **353**: 1877–81.

Olasehinde, G.I., Yah, C.S., Singh, R., Ojurongbe, O.O., Ajayi, A.A., Valecha, N., Abolaji, A.O. and Adeyeba, A.O. (2012). Genetic diversity of *Plasmodium falciparum* field isolates from southwestern Nigeria. *African Health Science*, **12**: 355–361.

Olasehinde, G.I., Ojurongbe, O., Adeyeba, A., Fagade, O., Valecha, N., Ayanda, I.O., Ajayi, A.A. and Egwari, L.O. (2014). In vitro studies on the sensitivity pattern of *Plasmodium falciparum* to anti-malarial drugs and local herbal extracts. *Malaria Journal*, **13**: 63.

Olotu, A., Fegan, G., Williams, T.N., Sasi, P. and Ogada, E. (2010). Defining clinical malaria: the specificity and incidence of endpoints from active and passive surveillance of children in rural Kenya. *PLoS ONE*, **5**: e15569.

Omole, M.K and Onademuren, O.T (2010). A survey of antimalarial drug use practices among

urban dwellers in Abeokuta, Nigeria. *African Journal of Biomedical Research*, **13**: 1–7

Onyabe, D.Y., Vajime, C.G., Nock, I.H., Ndams, I.S., Akpa, A.U., Alaribe, A.A. and Conn, J.E (2003). The distribution of M and S molecular forms of *Anopheles gambiae* in Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **97**: 605–608.

Orozco-ter Wengel, P., Kapun, M., Nolte, V., Kofler, R., Flatt, T. and Schlotterer, C. (2012). Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogenous trajectories of selected alleles. *Molecular Ecology*, **21** (20): 4931-4941.

Oquendo, P., Hundt, E., Lawler, J. and Seed, B. (1989). CD36 directly mediates cytoadherence of *Plasmodium falciparum* parasitized erythrocytes. *Cell*, **58**: 95-101.

Oyedeji, S.I., Awobode, H.O. and Kun, J. (2013). Limited genetic diversity and low multiplicity of *Plasmodium falciparum* infections in children with severe malaria in Lafia, North-Central Nigeria. *Journal of Experimental and Clinical Medicine*, **5**: 143–147.

Paget-McNicol, S., Gatton, M., Hastings, I. and Saul, A. (2002). The *Plasmodium falciparum* var gene switching rate, switching mechanism and patterns of parasite recrudescence described by mathematical modeling. *Parasitology*, **124**: 225-235.

Pagola, S., Stephens, P.W., Bohle, D.S., Kosar, A.D. and Madsen, S.K. (2000). The structure of malaria pigment β-haematin. *Nature*, **404**: 307-310.

Palacpac, N.M.Q., Ntege, E., Yeka, A., Balikagala, B., Suzuki, N., Shirai, H., Yagi, M., Ito, K., Fukushima, W., Hirota, Y., Nsereko, C., Okada, T., Kanoi, B.N. and Horii, T. (2013) Phase 1b randomized trial and follow-up study in Uganda of the blood-stage malaria vaccine candidate BK-SE36. *PLoS ONE*, DOI: 10.1371/journal.pone.0064073.

Pattanasin, S., Proux, S. and Chompasuk, D. (2003). Evaluation of a new *Plasmodium falciparum* lactate dehydrogenase assay (OptiMAL-IT) for the detection of malaria. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, **97**: 672-674.

Peakall, R and Smouse, P.E (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology*, **6**: 288–295.

Pespeni, M.H., Garfield, D.A., Manier, M.K. and Palumbi, S.R. (2012). Genome-wide polymorphisms show unexpected targets of natural selection. *Proceedings of the Royal Biological Sciences*, **279**: 1412-1420.

Phyo, A.P., Nkhoma, K., Stepniewska, E.A., Ashley, S., Nair, R., McGready, C and ler Moo, D (2012). Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study. *Lancet*, **379**: 1960-1966.

Polley, S.D., Tetteh, K.K., Lloyd, J.M., Akpoghenet, O.J., Greenwood, B.M., Bojang, K.A. and Conway, D.J. (2007). *Plasmodium falciparum* merozoite surface protein 3 is a target of allele-specific immunity and alleles are maintained by natural selection. *Journal of Infectious Diseases*, **195**: 279–287.

Poon, L.L., Wong, B.W., Ma, E.H., Chan, K.H., Chow, L.M., Abeyewickreme, W., Tangpukdee, N., Yuen, K.Y., Guan, Y., Looareesuwan, S. and Peiris, J.S. (2006). Sensitive and inexpensive molecular test for falciparum malaria: detecting *Plasmodium falciparum* DNA directly from heat-treated blood by loop-mediated isothermal amplification. *Clinical Chemistry*, **52**: 303-306.

Putaporntip, C., Jongwutiwes, S., Sakihama, N., Ferreira, M.U., Kho, W.G., Kaneko, A., Hattori,

T. and Tanabe, K. (2002). Mosaic organization and heterogeneity in frequency of allelic recombination of the *Plasmodium vivax* merozoite surface protein-1 locus. *Proceedings of the National Academy of Science USA*, **99**: 16348-16353.

Ranford-Cartwright, L., Johnston, K., Abdel-Muhsin, A., Khan, B. and Babiker, H. (2002). Critical comparison of molecular genotyping methods for detection of drug-resistant *Plasmodium falciparum. Transactions of the Royal Society of Tropical Medicine and Hygiene*, **96**: 568-572.

Redd, S.C., Bloland, P.B., Kazembe, P.N., Patrick, E., Tembenu, R. and Campbell, C.C. (1992). Usefulness of clinical case-definitions in guiding therapy for African children with malaria or pneumonia. *Lancet*, **340**: 1140-1143.

Rich, S.M. and Ayala, F.J. (2000). The origin of antigenic diversity in *Plasmodium falciparum*. *Parasitology Today*, **16**: 390-396.

Ridley, R.G. (2002). Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature*, **415**(6872): 686-693.

Rogerson, S.J., Chaiyaroj, S.C., Ng, K., Reeder, J.C. and Brown, G.V. (1995). Chondroitin sulfate A is a cell surface receptor for *Plasmodium falciparum*-infected erythrocytes. *Journal of Experimental Medicine*, **182**: 15-20.

Rozas, J., Sanchez-Delbarrio, J.C., Messegeur, X. and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**: 2496-2497.

Sa, J.M., Twu, O., Hayton, K. and Reyes, S. (2009) Geographic patterns of *Plasmodium*

falciparum drug resistance distinguished by differential responses to amodiaquine and chloroquine. *Proceedings of the National Academy of Science, USA*, **106**: 18883-18889.

Sagara, I., Dicko, A., Ellis, R.D., Fay, M.P., Diawara, S.I., Assadou, M.H., Sissoko, M.S., Kone, M., Diallo, A.I., Saye, R., Guindo, M.A.,Kante, O., Niambele, M.B., Miura, K., Mullen, G.E., Perce, M., Martin, L.B., Dolo, A., Diallo, D.A., Doumbo, O.K., Miller, L.H. and Saul, A. (2009). A randomized controlled phase 2 trial of the blood stage AMA1-C1/Alhydrogoel malaria vaccine in children in Mali. *Vaccine*, **27**(23): 3090-3098.

Sallusto, F., Lanzavecchia, A., Araki, K. and Ahmed, R. (2010). From vaccines to memory and back. *Immunity*, **33**(4): 451-463.

Saralamba, S., Pan-Ngum, W., Maude, R.J., Lee, S.J., Tarning, J., Lindegardh, N., Chotivanich, K., Nosten, F., Day, N.P.J., Socheat, D., White, N.J. and White, L.J. (2010). Intrahost modeling of artemisinin resistance in *Plasmodium falciparum*. *Proceedings of the National Academy of Science USA*, **108** (1): 397-402.

Schijns, V.E. and Lavelle, E.C. (2011). Trends in vaccine adjuvants. *Expert Review of Vaccines*, **10** (4): 539-550.

Schofield, L., Villaquiran, J. and Ferreira, A. (1987). Gamma interferon, CD8+ T cells and antibodies required for immunity to malaria sprozoites. *Nature*, **330**: 664-666.

Schultz, L., Wapling, J., Mueller, I., Ntsuke, P.O., Senn, N., Nale, J., Kiniboro, B., Buckee, C.O., Tavul, L., Siba, P.M., Reeder, J.C. and Barry, A.E (2010). Multilocus haplotypes reveal variable levels of diversity and population structure of *Plasmodium falciparum* in Papua New Guinea, a region of intense perennial transmission. *Malaria Journal*, **9**:336.

Singh, B., Sung, L.K., Matusop, A., Radhakrishnan, A., Shamsul, S.S. and Cox-Singh, J., (2004). A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *Lancet*, **363**: 1017–1024.

Smith, J. D., Subramanian, G., Gamain, B., Baruch, D.I. and Miller, L.H. (2000). Classification of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family. *Molecular Biochemical Parasitology*, **110** (2):293-310.

Snounou, G., Viriyakosol, S., Zhu, X.P., and Jarra, W (1998). High sensitivity of detection of malaria parasite by the use of nested polymerase chain reaction. *Molecular and Biochemical Parasitology*, **61**: 315-320.

Sowa, K.M., Cavanagh, D.R., Creasey, A.M., Raats, J., McBride, J., Sauerwein, R., Roeffen, W.F. and Arnot, D.E. (2001). Isolation of a monoclonal antibody forms a malaria patient-derived phage display library recognizing the Block 2 region of *Plasmodium falciparum* merozoite surface protein-1. *Molecular and Biochemical Parasitology*, **112**(1): 143-147.

Souza, V., Nguyen, T.T., Hudson, R.R., Pinero, D. and Lenski, R.E. (1993). Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: evidence for sex? *Proceedings of the National Academy of Science*, USA, **89**:8389–8393.

Sowa, K.M., Cavanagh, D.R., Creasey, A.M., Raats, J., McBride, J., Sauerwin, R., Roeffen, W.F., Arnot, D.E. (2001). Isolation of a monoclonal antibody from a malaria patient-derived phage display library recognizing the block 2 region of *Plasmodium falciparum* merozoite surface protein-1. *Molecular and Biochemical Parasitology*, **112**: 143-147.

Su, X., Heatwole, V.M., Wertheimer, S.P., Guinet, F., Herrfeldt, J.A., Peterson, D.S., Ravetch,

J.A.and Wellems. T.E. (1995). The Large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell*, **82**: 89-100.

Su, X., Kirkman, L.A., Fujioka, H. and Wellems, T.E. (1997). Complex polymorphisms in an approximately 330 kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. *Cell*, **91**: 593-603.

Su, X., Ferdig, M.T., Huang, Y., Huynh, C.Q., Liu, A., You, J., Wootton, J.C. and Wellems, T.E (1999). A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. *Science*, **286**: 1351–1353.

Sutherland, C.J. and Hallett, R. (2009). Detecting malaria parasites outside the blood. *Journal of Infectious Diseases*, **199**(11): 1561-1563.

Sutherland C., Tanomsing N., Nolder D., Oguike M., Jennison C., Pukrittayakamee S., Dolecek C., Hien T.T., do Rosário V.E., Arez A.P., Pinto J., Michon P., Escalante A., Nosten F., Burke M., Lee R., Blaze M., Otto T.D., Barnwell J., Pain A., Williams J., White N.J., Day N.P.J., Snounou G., Lockhart P.J., Chiodini P.L., Imwong M. and Polley S.D. (2010). Two non-recombining sympatric forms of the human malaria parasite *Plasmodium ovale* occur globally. *Journal of Infectious Diseases*, **201**:1544–1550

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**: 585–595.

Takahata, N. and Satta, Y. (1998). Footprints of intragenic recombination at HLA loci. *Immunogenetics*, **47**: 430-441.

Takala, S.L., Coulibaly, D., Thera, M.A., Batchelor, A.H., Cummings, M.P., Escalante, A.A., Ouattara, A., Traore, K., Niangaly, A., Djimde, A.A., Doumbo, O.K. and Plowe, C.V. (2007). Extreme polymorphism in a vaccine antigen and risk of clinical malaria: Implications for vaccine development. *Science and Translational Medicine*, **1**(2): 2ra5.

Takala, S.L. and Plowe, C.V. (2009). Genetic diversity and malaria vaccine design, testing and efficacy: Preventing and overcoming "vaccine resistant malaria". *Parasite Immunology*, **31**: 560-573.

Tami, A., Ord, R., Targett, G.A. and Sutherland, C.J. (2003). Sympatric *Plasmodium falciparum* isolates from Venezuela have structured var gene repertoires. *Malaria Journal*, **2:** 7.

Tan, J.C., Patel, J.J., Tan, A., Blain, J.C., Albert, T.J., Lobo, N.F. and Ferdig, M.T. (2009). Optimizing comparative genomic hybridization probes for genotyping and SNP detection in *Plasmodiun falciparum. Genomics*, **93**: 543-550.

Tanabe, K., Mackay, M., Goman, M. and Scaife, J.G. (1987). Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. *Journal of Molecular Biology*, **195**: 273-287.

Taylor, R.R., Smith, D.B., Robinson, V.J., McBride, J.S. and Riley, E.M. (1995). Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of immunoglobulin G3 subclass. *Infectious Immunology*, **63**: 4382-4388.

Tebo, A.E., Kremsner, P.G. and Luty, A.J. (2001). *Plasmodium falciparum*: A major role for IgG3 in antibody-dependent monocyte mediated cellular inhibition of parasite growth in vitro. *Experimental Parasitology*, **98**: 20-28.

Temu, E.A., Kimani, I., Tuno, N. and Kawada, H. (2006). Monitoring chloroquine resistance using *Plasmodium falciparum* parasites isolated from wild mosquitoes in Tanzania. *American Journal of Tropical Medicine and Hygiene*, **75**: 1182-1187.

Tetteh, K.K., Stewart, L.B., Ochola, L.I., Amambua-Ngwa, A., Thomas, A.W., Marsh, K., Weedall, G.D and Conway, D.J. (2009). Prospective identification of malaria parasite genes under balancing selection. *PLoS One*. **4:**e5568.

Thekisoe, O.M., Inoue, N., Kuboki, N., Tuntasuvan, D., Bunnoy, W., Borisutsuwan, S., Igarashi, I. and Sugimoto, C. (2005). Evaluation of loop-mediated isothermal amplification (LAMP), PCR and parasitological tests for detection of *Trypanosoma evansi* in experimentally infected pigs. *Veterinary Parasitology*, **130**: 327-330.

Tinto, H., D'Allessandro, U., Sorgho, H., Tahita, M.C., Kabore, W., Kiemde, F., Lompo, P., Ouedraogo, S., Derre, K. Kaslow, D., Leboullex, D., Savarese, B. and Schellenberg, D. (2015). Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase , individually randomised, controlled trial. *Lancet*, **386**: 31-45.

Tong-Soo, K., Hyung-Hwan, K., Sun-Sim, L., Byoung-Kuk, N., Khin, L., Shin-Hyeong, C., Yoon-Joong, K.,Do-Kyung, K., Youngjoo, S., Kim, H and Hyeong-Woo, L. (2010). Prevalence of *Plasmodium vivax* VK210 and VK247 subtype in Myanmar. *Malaria Journal*, **9**: 195.

Treutiger, C.J., Hedlund, I., Helmby, H., Carlson, J., Jepson, A., Twumasi, P., Kwiatkowski, D., Greenwood, B.M. and Wahlgren, M. (1992). Rosette formation in *Plasmodium falciparum* isolates and anti-rosette activity of sera from Gambians with cerebral or

uncomplicated malaria. *American Journal of Tropical Medicine and Hygiene*, **46** (5):503-510.

Triglia, T., Wang, P., Sims, P.F.G., Hyde, J.E. and Cowman, A.F. (1998). Allelic exchange at the endogenous genomic locus in *Plasmodium falciparum* proves the role of dihdropteroate synthase in sulfadoxine-resistant malaria. *European Molecular Biology Organization Journal*, **17**: 3807-3815.

Van Tyne, D., Park, D.J., Schaffner, S.F., Neafsey, D.E. and Angelino, E. (2011). Identification and functional validation of the novel antimalarial resistance locus PF10_0355 in *Plasmodium falciparum. PLoS Genetics*, **7**: e10011383.

Vazquez-Macias, A., Martinez-Cruz, P., Castaneda-Patlan, M.C., Scheidig, C., Gysin, J. and Scherf, A. (2002) A distinct 5' flanking var gene region regulates *Plasmodium falciparum* variant erythrocyte surface antigen expression in placental malaria. *Molecular Microbiology*, **45**: 155–167.

Venkatessan, M., Amaratunga, C., Campino, S., Auburn, S., Koch, O., Lim, P., Uk, S., Socheat, D., Kwiatkowski, D.P., Fairhurst, R.M. and Plowe, C.V. (2012). Using CF11 cellulose columns to inexpensively and effectively remove human DNA from *Plasmodium falciparum*-infected whole blood samples. *Malaria Journal*, **11**: 41

Veron, V., Legrand, E., Yrinesi, J., Volney, B., Simon, S. and Carme, B. (2009). Genetic diversity of msp3α and msp1_b5 markers of *Plasmodium vivax* in French Guiana. *Malaria Journal*, **8**: 40.

Verra, F., Chokejindachai, W., Weedall, G.D., Polley, S. D., Mwangi, T.W., Marsh, K. and

Conway, D.J. (2006). Contrasting signatures of selection on the *Plasmodium falciparum* erythrocyte binding antigen gene family. *Molecular and Biochemical Parasitology*. **149**: 182–190.

Verstrepen, K.J., Jansen, A., Lewitter, F. and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. *Nature Genetics*, **37**: 986-990.

Vogel, G. (2013). The Forgotten malaria. *Science*, **2013**: 684-687.

Vogt, A.M., Barragan, A., Chen, Q., Kironde, F., Spillman, D. and Wahlgreen, M. (2003). Heparan sulfate on endothelial cells mediates the binding of *Plasmodium falciparum*-infected erythrocytes via the DBL1alpha domain of PfEMP1. *Blood*, **101**: 2405-2411.

Voight, B.F., Kudaravalli, S., Wen, S. and Pritchard, J.K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*: 10.1371

Volkman, S.K., Sabeti, P.C. and DeCaprio D. (2007). A genome-wide map of diversity in *Plasmodium falciparum. Nature Genetics*, **39**: 113–119.

Voss, T. S., Healer, J., Marty, A.J., Duffy, M.F., Thompson, J.K., Beeson, J.G., Reeder, J.C., Crabb, B.S. and Cowman, A.F. (2006). A var gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature,* **439** (7079):1004-8.

Vythiligan, C.H., Asmad M. and Chan S.T (2008). Natural transmission of *Plasmodium knowlesi* to humans by *Anopheles latens* in Sarawak, Malaysia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **100**:1087–1088.

Wang, P., Sims, P.F. and Hyde, J.E. (1997). A modified in vitro sulfadoxine susceptibility assay

for *Plasmodium falciparum* suitable for investigating Fansidar resistance. *Parasitology*, **115**: 223-230.

Wargo, A.R., Huijben, S., de Roode, J.C., Sheperd, J. and Read, A.F. (2007). Competitive release abd facilitation of drug-resistant parasites after therapeutic chemotherapy in a rodent malaria model. *Proceedings of the National Academy of Science*, **104**: 19914-19919.

Warhurst, D.C and Williams, J.C (1996). ACP broadsheet No. 148. Laboratory diagnosis of malaria. *Journal of Clinical Pathology*, **49**: 533-538.

Watanabe, J., Suzuki, Y., Sasaki, M. and Sugano, S. (2004). Full-malaria-2004: An enlarged database for comparative studies of full-length cDNAs of malaria parasites, *Plasmodium* species. *Nucleic Acid Research*, **32**: D334-D338.

Weber, J.L. (1988). Interspersed repetitive DNA from *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*, **29**:117-124.

Weedall, G.D., Preston, B.M., Thomas, A.W., Sutherland, C.J. and Conway, D.J. (2007). Differential evidence of natural selection on two leading sporozoite stage malaria vaccine candidate antigens. *International Journal of Parasitology*, **37**: 77–85.

Westling, J., Yowell, C. A., Majer, P., Erickson, J. W., Dame, J. B. and Dunn, B. M. (1997). *Plasmodium falciparum*, *P. vivax*, and *P. malariae*: A Comparison of the Active Site Properties of Plasmepsins Cloned and Expressed from Three Different Species of the Malaria Parasite. *Experimental Parasitology*, **87** (3): 185–193.

Wiseman, V., Hawley, W.A., ter Kuile, F.O., Phillips-Howard , P.A., Vulule, J.M. and Nahlen,

B.L. (2003). The cost-effectiveness of permethrin-treated bed nets in an area of intense malaria transmission in western Kenya. *The American Journal of Tropical Medicine and hygiene*, **68**: 161-167.

World Health Organization (2003). Assessment and monitoring of antimalarial drug efficacy for the treatment of uncomplicated falciparum malaria. World Health Organization: Geneva, WHO/HTM/RBM/2003.5

World Health Organization (2006). Use of indoor residual spraying for scaling up global malaria control and elimination. WHO/HTM/MAL/2006.112, 10.

World Health Organization (2008). World Malaria Report 2008. WHO/HTM/GMP/2008

World Health Organization (2010). Guidelines for the Treatment of Malaria. 2nd edition. WHO, Geneva.

World Health Organization (2012). Malaria Vaccine Rainbow Tables. *WHO*, Geneva.

World Health Organization (2014). World Malaria Report 2014. *WHO*, Geneva

World Health Organization (2015). World Malaria Report 2015. *WHO*, Geneva

Wootton, J.C., Feng, X, Ferdig, M.T., Cooper, R.A., Mu, J., Baruch, D.I., Magill, A.J. and Su, X.Z. (2002). Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum. Nature*, **418**: 320–323.

Wright, P.E. and Dyson, H.J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology*, **293**: 321-331.

Zhong, D., Afrane, Y., Githeko, A. and Cui, L. (2008). Molecular epidemiology of drug-resistant

malaria in western Kenya highlands. *BMC Infectious Diseases*, **8**:105.

Zuk, M. and McKean, K.A. (1996). Sex differences in parasite infections: patterns and processes. *International Journal of Parasitology*, **26**: 1009-1023.

Appendix 1: Nigerian Institute of Medical Research Review Board Approval Letter

INSTITUTIONAL REVIEW BOARD

IRB-NIMR

NIGERIAN INSTITUTE OF MEDICAL RESEARCH

6, Edmond Crescent Off Murtala Muhammed Way, P.M.B. 2013 Yaba, Lagos.
Tel: 01-4823123, 01-7744723, 08050254484, 08033460947 Fax: 01-4823123, 234-1-3425171
E-mail: nimr_irb@yahoo.com *Website:* www.nimr-nig.org
Secretariat: Room 207, Biochemistry Division, Research Block, NIMR

28th February, 2013

PROJECT TITLE: **GENETIC DYNAMICS AND POPULATION STRUCTURE OF *PLASMODIUM FALCIPARUM* IN LAGOS AND EKITI STATES, SOUTH-WESTERN NIGERIA.**

PROJECT №: **IRB/12/209**

APPROVAL LETTER

The above named proposal has been adequately reviewed; the protocol and safety guidelines satisfy the conditions of NIMR-IRB, policies regarding experiments that use human subjects.

Therefore the study under its reviewed state is hereby approved by Institutional Review Board, NIMR.

PROF. F. E. OKONOFUA
Name of IRB Chairman

.....
Signature of IRB Chairman & Date

MRS. O. A. NWOGBE
Name of IRB Secretary

.....
Signature of IRB Secretary & Date

This approval is given with the investigator's Declaration as stated below;
By signing below I agree/certify that:

1. I have reviewed this protocol submission in its entirety and that I am fully cognizant of, and in agreement with, all submitted statements.
2. I will conduct this research study in strict accordance with all submitted statements except where a change may be necessary to eliminate an apparent immediate hazard to a given research subject.
   - I will notify the IRB promptly of any change in the research procedures necessitated in the interest of the safety of a given research subject.
   - I will request and obtain IRB approval of any proposed modification to the research protocol or informed consent document(s) prior to implementing such modifications.

Appendix 1 cont'd: Ethical approval letter (backpage)

3. I will ensure that all co-investigators and other personnel assisting in the conduct of this research study have been provided a copy of the entire current version of the research protocol and are fully informed of the current (a) study procedures (including procedure modifications); (b) informed consent requirements and process; (c) potential risks associated with the study participation and the steps to be taken to prevent or minimize these potential risks; (d) adverse event reporting requirements; (e) data and record-keeping; and (f) the current IRB approval status of the research study.

4. I will respond promptly to all requests for information or materials solicited by the IRB or IRB Office.

5. I will submit the research study in a timely manner for IRB renewal approval.

6. I will not enroll any individual into this research study until such time that I obtain his/her written informed consent, or, if applicable, the written informed consent of his /her authorized representative (i.e., unless the IRB has granted a waiver of the requirement to obtain written informed consent).

7. I will employ and oversee an informed consent process that ensures that potential research subjects understand fully the purpose of the research study, the nature of the research procedures they are being asked to undergo, the potential risks of these research procedures, and their rights as a research study volunteer.

8. I will ensure that research subjects are kept fully informed of any new information that may affect their willingness to continue to participate in the research study.

9. I will maintain adequate, current, and accurate records of research data, outcomes, and adverse events to permit an ongoing assessment of the risks/benefit ratio of research study participation.

10. I am cognizant of, and will comply with, current federal regulations and IRB requirements governing human subject research including adverse event reporting requirements.

11. I will make a reasonable effort to ensure that subjects who have suffered an adverse event associated with research participation receive adequate care to correct or alleviate the consequences of the adverse event to the extent possible.

12. I will ensure that the conduct of this research study adheres to Good Clinical Practice guidelines

MR. OYEBOLA KOLAPO
Principal Investigator's Name

mykolapo 28/02/2013
.....
Principal Investigator's Signature and Date

Appendix 2: Sequences of the primers used to amplify MSP-1 and MSP-2 genes of *P. falciparum* isolates

Locus	Primer	Primer sequence
Primary PCR		
msp-1	*MSP-1*-P1	5′ -CACATGAAAGTTATCAAGAACTTGTC-3′
	MSP-1-P2	5′ -GTACGTCTAATTCATTTGCAC-3′
msp-2	*msp*2-1	5′ -ATGAAGGTAATTAAAACATTGTCTATTATA-3′
	*msp*2-4	5′ -ATATGGCAAAAGATAAAACAAGTG-3′
Secondary PCR		
msp-1	K1-K1	5′ -GAAATTACTACAAAAGGTGCAAGTG-3′
	K1-K2	5′ -AGATGAAGTATTTGAACGAGGTAAAGTG-3′
	MAD20-M1	5′ -GAACAAGTCGAACAGCTGTTA-3′
	MAD20-M2	5′ -TGAATTATCTGAAGGATTTGTACGTCTTGA-3′
	R033-R1	5′ -GCAAATACTCAAGTTGTTGCAAAGC-3′
	R033-R2	5′ -AGGATTTGCAGCACCTGGAGATCT-3′
msp-2	3D7-A1	5′ -GCAGAAAGTAAGCCTTCTACTGGTGCT-3′
	3D7-A2	5′ -GATTTGTTTCGGCATTATTATGA-3′
	FC27-B1	5′ -GCAAATGAAGGTTCTAATACTAATAG-3′
	FC27-B2	5′ -GCTTTGGGTCCTTCTTCAGTTGATTC-3′

Appendix 3: Sequences of the primers used to amplify microsatellite loci

Name	Label	Sequence	Chromosome
TA1-3(F)	Blue	CTACATGCCTAATGAGCA	6
TA1-R		TTTTATCTTCATCCCCAC	
TA1-F		CCGTCATAAGTGCAGAGC	
TA40-3(R)	Yellow	AGCCACTGTATCCAGCCA	10
TA40-F		AAGGGATTGCTGCAAGGT	
TA40-R(2)			
Polya-R	Green	ATCAGATAATTGTTGGTA	4
Polya-F		AAAATATAGACGAACAGA	
Polya-3(IR)		GAAATTATAACTCTACCA	
TAA60-F	Green	CTCAAAGAAAAATAATTCA	13
TAA60-R		AAAAAGGAGGATAAATACAT	
TAA60-3(IF)		TAGTAACGATGTTGACAA	
ARA2-3(F)	Yellow	GTACATATGAATCACCAA	11
ARA2-R		GCTTTGAGTATTATTAATA	
ARA2-F		GAATAAACAAAGTATTGCT	
Pfg377-3(R)	Yellow	TTATGTTGGTACCGTGTA	12
Pfg377-F		GATCTCAACGGAAATTAT	
Pfg377-R		TTATCCCTACGATTAACA	
PfPK2-3(R)	Yellow	CCTCAGACTGAAATGCAT	12
PfPK2-F		CTTTCATCGATACTACGA	
PfPK2-R		AAAGAAGGAACAAGCAGA	
TAA87-3(F)	Green	ATGGGTTAAATGAGGTACA	6
TAA87-R		ACATGTTCATATTACTCAC	
TAA87-F		AATGGCAACACCATTCAAC	

Appendix 3 (Cont'd): Sequences of the primers used to amplify microsatellite loci

TAA109-3(F)	Blue	TAGGGAACATCATAAGGAT	6
TAA109-R		CCTATACCAAACATGCTAAA	
TAA109-F		GGTTAAATCAGGACAACAT	
TAA81-3(F)	Blue	GAAGAAATAAGGGAAGGT	5
TAA81-R		TTTCACACAACACAGGATT	
TAA81-F		TGGACAAATGGGAAAGGATA	
TAA42-3(F)	Green	ACAAAAGGGTGGTGATTCT	5
TAA42-R		GTATTATTACTACTACTAAAG	
TAA42-F		TAGAAACAGGAATGATACG	
2490-3(R)	Blue	ATGATGTGCAGATGACGA	
2490-F		TTCTAAATAGATCCAAAG	
2490-R		TAGAATTATTGAATGCAC	

Appendix 4: Allele frequencies at the 10 microsatellite loci examined in each of the *P. falciparum* isolates from three localities

Locus	Allele	Lekki (Frequency)	Badagry (Frequency)	Aramoko-Ekiti (Frequency)
Poly α	107	0	0.06	0.07
	110	0	0	0.02
	113	0.1	0.06	0.02
	116	0	0.24	0.12
	119	0.3	0.24	0.24
	122	0.5	0.12	0.27
	125	0.1	0.06	0.10
	128	0	0.12	0.10
	131	0	0	0.02
	134	0	0.06	0.02
	172	0	0.06	0
	Total	**30**	**51**	**82**
PFPK2	154	0	0.07	0
	157	0	0.07	0
	160	0.13	0.13	0.05
	163	0.25	0.27	0.23
	166	0	0.27	0.23
	169	0.25	0.07	0.29
	172	0.38	0.0	0.05
	175	0	0	0.08
	178	0	0	0.05
	181	0	0	0.05
	184	0.13	0.07	0.03
	187	0	0	0.05
	278	0	0	0.03
	Total	**24**	**45**	**76**
TA81	120	0	0	0.01
	141	0	0	0.01
	144	0	0	0.01
	147	0	0	0.01
	150	0	0	0.03
	153	0	0.07	0.03
	156	0.11	0.14	0.28
	159	0.22	0.14	0.22
	162	0.22	0.14	0.14
	165	0	0	0.08
	168	0.22	0.21	0.06
	171	0.11	0.14	0.03

Appendix 4 (cont'd): Allele frequencies at the 10 microsatellite loci examined in each of the *P. falciparum* isolates from three localities

	174	0.11	0	0.08
	177	0	0.14	0
	180	0	0	0.11
	Total	**27**	**42**	**72**
ARA 2	50	0	0.06	0.04
	53	0	0.12	0
	65	0.63	0.12	0.19
	68	0.25	0.18	0.26
	71	0.125	0	0.13
	74	0	0	0.02
	77	0	0	0.02
	80	0	0	0.04
	86	0	0.12	0.02
	89	0	0.06	0
	110	0	0.35	0.28
	Total	**24**	**51**	**94**
TA 40	63	0	0.06	0
	81	0	0	0.02
	90	0	0	0.02
	96	0	0.13	0.02
	99	0.11	0.06	0.17
	102	0.11	0.31	0.14
	105	0.22	0.19	0.21
	108	0.11	0	0.14
	111	0.44	0	0.12
	114	0	0.06	0.10
	117	0	0.06	0
	120	0	0.13	0.05
	Total	**27**	**48**	**84**
TA 42	161	0.29	0.17	0.13
	164	0.14	0.17	0.09
	167	0	0.33	0.09
	170	0.29	0	0.30
	173	0	0.33	0.04
	176	0.29	0	0.04
	179	0	0	0.04
	182	0	0	0.09
	185	0	0	0.04
	188	0	0	0.09
	191	0	0	0.04
	Total	**21**	**18**	**46**
2490	62	0.83	0.83	0.3
	186	0.17	0.17	0.6

Appendix 4 (cont'd): Allele frequencies at the 10 microsatellite loci examined in each of the *P. falciparum* isolates from three localities

	201	0	0	0.05
	246	0	0	0.05
	Total	**18**	**18**	**40**
TA 60	67	0	0.19	0.29
	94	0	0.06	0.05
	97	0.38	0.25	0.19
	100	0.50	0.50	0.45
	103	0.13	0	0
	106	0	0	0.02
	Total	**24**	**48**	**84**
TA 109	156	0.17	0	0
	159	0	0	0
	162	0	0.18	0.18
	165	0	0.12	0.18
	171	0	0.06	0
	174	0.33	0.12	0.18
	177	0.33	0.12	0.16
	180	0	0	0.07
	186	0	0	0.02
	189	0.17	0	0
	198	0	0.41	0.22
	Total	**18**	**51**	**90**
PFG377	76	0.25	0.27	0.22
	79	0	0.33	0.22
	82	0	0	0.06
	85	0.38	0.27	0.25
	88	0.13	0.13	0.25
	94	0.13	0	0
	97	0.13	0	0
	Total	**24**	**45**	**72**

Appendix 5: Amplification primers and double-labelled hydrolysis probes for parasite (PgMET) and human (HumTuBB) targets

Primer/probe	5′ Fluorophore	Sequence	3′Quencher
PgMET_F1	FAM	5′-TGAAAGCAGCGTAGCTCAGA	BHQ2
PgMET-R2		5′-CGCGTGGTTTCGATCCACG	
PgMET-pB		5′-GGGGCTCATAACCCCCAGGA	
HumTuBB-F2	JOE	5′-AAGGAGGTCGATGAGCAGAT	BHQ2
HumTuBB-R2		5′-GCTGTCTTGACATTGTTGGG	
HumTuBB-Joe		5′-TTAACGTGCAGAACAAGAACAGCAGCT	

Appendix 6: Reaction recipe for RT-PCR detection of human ribonuclease P

Reaction components	Quantity (ul)	
	X 1	X 100
Rnase primer probe (FAM-labeled)	1.25	125
RT-PCR enzyme mix	12.5	1250
Molecular grade Rnase-free water	9.25	925
Microbiome-enriched DNA	2	

Appendix 7: Recipe for end-repair of sheared DNA

Component	10x reaction (ul)
DNA (100ng sheared DNA)	15
RNase-free water	55
End-repair Buffer, 10x	25
End-repair Enzyme mix	20
Total reaction volume	115

Appendix 8: Recipe for adapter ligation of end-repaired DNA

Component	Singleplex adapter mix- 10x reaction volume (ul)
End-repaired DNA	25
Ligation buffer 2x	400
Adapter mix	100
Ligation and Nick repair mix	40
dNTP mix (10mM)	10
RNase-free water	variable

Appendix 9: PCR quantitation of sample DNA

Sample	Cq Value	Starting quantity	500,000/µl normalisation	Amount in the DNA pool
D06	16.41	6657120.58632	0.07511	4µl
A07	17.95	2039432.55451	0.24517	4µl
B04	18.28	1691523.77919	0.29559	4µl
H08	18.82	1053613.10928	0.47456	4µl
H06	19.06	873633.08419	0.57232	4µl
H09	19.35	697562.62731	0.71678	4µl
B06	20.02	453216.39395	1.10323	4µl
A06	20.09	431208.28965	1.15953	4µl
G04	20.30	339188.68302	1.47411	4µl
D05	20.64	260505.66903	1.91934	4µl
C10	21.03	211921.79911	2.35936	10µl
H12	20.94	207002.09110	2.41543	10µl
A06	21.05	191054.41090	2.61706	10µl
F07	21.06	188552.07833	2.65179	10µl
H05	21.64	121660.37072	4.10980	10µl
C08	21.71	115222.08244	4.33945	10µl
C06	21.89	99985.88121	5.00071	10µl
A11	22.17	89370.39951	5.59469	10µl
D04	22.20	87678.36913	5.70266	10µl
E08	22.26	75869.72106	6.59024	10µl
D10	22.32	72396.41152	6.90642	10µl
E06	22.32	72003.67082	6.94409	10µl
A08	22.55	67418.85451	7.41632	10µl
C12	22.61	64399.50014	7.76404	10µl
C05	22.63	63343.09941	7.89352	10µl
C04	22.65	62225.11898	8.03534	10µl
A07	22.73	58961.15103	8.48016	10µl
G06	22.61	57935.99753	8.63021	10µl
G11	22.66	55616.48871	8.99014	10µl
C09	22.86	53172.12409	9.40342	10µl
A09	22.90	51829.95701	9.64693	10µl
G10	22.91	46107.96929	10.84411	10µl
C11	23.09	44820.79053	11.15554	20µl
A12	23.11	43991.76316	11.36576	20µl
E09	22.99	43223.47576	11.56779	20µl
H07	23.05	41430.61171	12.06837	20µl
B10	23.23	40180.78150	12.44376	20µl
A10	23.13	38866.64236	12.86450	20µl
E07	23.32	33575.05058	14.89201	20µl

Appendix 9 (cont'd): PCR quantitation of sample DNA

G08	23.40	31612.48150	15.81654	20µl
A10	26.45	3559.50543	18905.54908	20µl
F12	26.32	3379.37660	18994.98556	20µl
D05	25.99	5043.11129	19241.54098	20µl
E10	25.83	4944.13773	19360.88668	20µl
A04	25.73	6108.50865	19431.51591	20µl
D08	25.67	5547.44838	19474.3957	20µl
H10	25.53	6188.19264	19583.39506	20µl
H11	25.38	6928.52316	19697.37294	20µl
C10	25.27	7535.58394	19782.94993	20µl
B08	25.26	8721.07813	19794.59077	20µl
A05	25.05	8920.98249	19957.17745	20µl
G12	25.03	9092.72130	19977.05537	20µl
E05	24.50	13615.91229	20407.46404	20µl
E04	24.47	13947.20169	20433.67589	20µl
C07	24.41	16509.84299	20480.51014	20µl
G07	24.27	16224.64691	20600.15192	20µl
B09	24.21	19318.52939	20656.71576	20µl
B10	24.08	18747.42124	20761.77709	20µl
A05	23.97	23024.81194	20857.17798	20µl
C06	23.86	25099.40240	20957.14926	20µl
B07	23.78	26645.76936	21026.99356	20µl
B05	23.65	26009.88769	21137.48973	20µl
A11	23.49	29423.29522	21282.5361	20µl
D11	23.48	29618.65506	21290.37662	20µl
B11	23.46	30230.40563	21314.63464	20µl

Appendix 10: Genes within the *Plasmodium falciparum* genome with positive Tajima's D values

Genes	Number of SNPS	Tajima's D values
PF3D7_1437000	12	0.00
PF3D7_1437400	3	0.01
PF3D7_1437500	86	0.01
PF3D7_1437600	12	0.02
PF3D7_1437700	17	0.02
PF3D7_1437900	28	0.03
PF3D7_1438000	8	0.03
PF3D7_1438400	5	0.05
PF3D7_1438600	17	0.05
PF3D7_1438700	32	0.05
PF3D7_1438800	16	0.07
PF3D7_1439100	13	0.09
PF3D7_1439200	3	0.10
PF3D7_1439300	28	0.10
PF3D7_1439400	11	0.10
PF3D7_1439500	7	0.11
PF3D7_1439700	13	0.11
PF3D7_1440000	17	0.14
PF3D7_1440100	9	0.16
PF3D7_1440200	6	0.16
PF3D7_1440400	6	0.16
PF3D7_1440500	16	0.20
PF3D7_1440600	15	0.22
PF3D7_1440700	5	0.23
PF3D7_1440800	4	0.23
PF3D7_1440900	21	0.25
PF3D7_1441000	7	0.25
PF3D7_1441300	3	0.26
PF3D7_1441500	22	0.26
PF3D7_1441600	86	0.28
PF3D7_1441700	11	0.32
PF3D7_1442100	14	0.33
PF3D7_1442200	27	0.33
PF3D7_1442300	50	0.35
PF3D7_1442400	23	0.35
PF3D7_1442600	5	0.35
PF3D7_1442700	114	0.38
PF3D7_1442800	3	0.39
PF3D7_1442900	94	0.41

Appendix 10 (cont'd): Genes within the *Plasmodium falciparum* genome with positive Tajima's D values

PF3D7_1443000	10	0.42
PF3D7_1443100	19	0.43
PF3D7_1443200	6	0.44
PF3D7_1443400	4	0.45
PF3D7_1443500	19	0.46
PF3D7_1443600	13	0.48
PF3D7_1443800	9	0.49
PF3D7_1443900	9	0.51
PF3D7_1444000	14	0.53
PF3D7_1444100	18	0.53
PF3D7_1444400	6	0.56
PF3D7_1444500	36	0.58
PF3D7_1444700	15	0.64
PF3D7_1445100	3	0.64
PF3D7_1445200	9	0.65
PF3D7_1445300	6	0.65
PF3D7_1445400	7	0.70
PF3D7_1445600	7	0.70
PF3D7_1445800	16	0.75
PF3D7_1445900	4	0.78
PF3D7_1446100	3	0.79
PF3D7_1446200	7	0.82
PF3D7_1446300	15	0.82
PF3D7_1446500	14	0.85
PF3D7_1446700	56	0.86
PF3D7_1446900	33	0.87
PF3D7_1447100	8	0.87
PF3D7_1447200	28	0.88
PF3D7_1447400	7	0.91
PF3D7_1447600	25	0.98
PF3D7_1447800	50	0.99
PF3D7_1447900	6	1.00
PF3D7_1448000	14	1.04
PF3D7_1448200	6	1.05
PF3D7_1448300	11	1.06
PF3D7_1448400	227	1.07
PF3D7_1448500	5	1.10
PF3D7_1448600	65	1.16
PF3D7_1448700	146	1.20
PF3D7_1448800	9	1.25
PF3D7_1449000	24	1.28
PF3D7_1449100	16	1.30
PF3D7_1449300	131	1.32
PF3D7_1449400	26	1.33
PF3D7_1449500	230	1.36

Appendix 10 (cont'd): Genes within the *Plasmodium falciparum* genome with positive Tajima's D values

PF3D7_1449600	70	1.45
PF3D7_1449700	14	1.48
PF3D7_1449900	4	1.58
PF3D7_1450100	17	1.65
PF3D7_1450300	3	1.67
PF3D7_1450400	14	1.86
PF3D7_1450500	21	2.03
PF3D7_1450600	19	2.09
PF3D7_1450700	11	2.11
PF3D7_1450800	18	2.34
PF3D7_1450900	25	2.46
PF3D7_1451000	84	2.68

Appendix 11: Published data on genetic diversity of *P. falciparum* in Lagos State

Asian Pacific Journal of Tropical Biomedicine

journal homepage: www.apjtb.com

Document heading doi:10.12980/APJTB.4.2014C1301

Genetic diversity and complexity of *Plasmodium falciparum* infections in Lagos, Nigeria

Muyiwa K Oyebola[1,2*], Emmanuel T Idowu[1], Yetunde A Olukosi[2], Bamidele A Iwalokun[2], Chimere O Agomo[2], Olusola O Ajibaye[2], Monday Tola[2], Adetoro O Otubanjo[1]

[1]*Parasitology and Bioinformatics, Faculty of Science University of Lagos, Nigeria*

[2]*Malaria Research Laboratory, Nigeria Institute of Medical Research, Lagos, Nigeria*

PEER REVIEW

Peer reviewer

Chien-Ming Shih, Professor, Department of Parasitology and Tropical Medicine, Graduate Institute of Pathology and Parasitology, National Defense Medical Center, Taipei 114, Taiwan.
Tel: +886-2-87923152 ext.18564
Fax: +886-2-87921821
E-mail: cmshih@ndmctsgh.edu.tw

Comments

This is an interesting investigation regarding the genetic diversity of *P. falciparum* populations within infected patients and results provide the existing genetic structure of *P. falciparum* in the country. The need for an additional investigation to understand the relationship between seasonality and diversity of *P. falciparum* populations is suggested.

Details on Page S90

ABSTRACT

Objective: To analyse the genetic diversity of *Plasmodium falciparum* (*P. falciparum*) using *msp*-1 and *msp*-2 as antigenic markers.

Methods: Parasite DNA was extracted from 100 blood samples collected from *P. falciparum*-positive patients confirmed by microscopy, and followed by PCR-genotyping targeting the *msp*-1 (block2) and *msp*-2 (block 3) allelic families.

Results: All the families of *msp*-1 (K1, MAD20 and R033) and *msp*-2 (FC27 and 3D7) locus were observed. Results revealed that K1 (60/100) was the most predominant genotype of *msp*-1 allelic family followed by the genotypes of MAD20 (50/100) and R033 (45/100). In the *msp*-2 locus, FC27 genotype (62/100) showed higher frequency than 3D7 genotype (55/100). The allelic families were detected either alone or in combination with other families. However, no R033/MAD20 combination was observed. Multiplicity of infection (MOI) with *msp*-1 was higher in the locality of Ikorodu (1.50) than in Lekki (1.39). However, MOI with *msp*-2 was lower in the locality of Ikorodu (1.14) than in Lekki (1.76). There was no significant difference in the mean MOI between the two study areas (P=0.427).

Conclusions: The observation of limited diversity of malaria parasites may imply that the use of antigenic markers as genotyping tools for distinguishing recrudescence and re-infections with *P. falciparum* during drug trials is subjective.

KEYWORDS

Diversity, Antigenic markers, Multiplicity of infections, Recrudescence, Drug trials

1. Introduction

Malaria still remains an important public health disease in the tropical parts of the world, especially in the African continent. In Nigeria, the disease is responsible for 60% outpatient visits to health facilities, 30% deaths of under-five childrens, 11% maternal mortality, and an estimated 300 000 deaths per year[1]. Contributing to the burgeoning burden of the disease is drug resistance that has crippled most antimalarial drugs[2-7]. In order to track the efficacy of existing antimalarial drugs, therapeutic efficacy trials are carried out over a follow-up period[8]. Since the probability of a newly infected patient possessing a parasite genotype identical to the former infection is low[9], molecular genotyping of pre-treatment (baseline) and recurrent infections enables the categorisation of recurrent parasites as recrudescent (*i.e.* true failure) or re-infection (*i.e.* successful treatment) either from pre-existing infection or

*Corresponding author: Muyiwa Kolapo Oyebola, Parasitology and Bioinformatics, Faculty of Science University of Lagos; Malaria Research Laboratory, Nigeria Institute of Medical Research, Lagos, Nigeria.
Tel: +2348034778549
E-mail: oyebolakolapo@yahoo.com

Foundation Project: partly supported by the Royal Society of Tropical Medicine and Hygiene Centenary Scholarship awarded to YAO in 2009.

Article history:
Received 4 Jan 2014
Received in revised form 12 Jan, 2nd revised form 19 Jan, 3rd revised form 6 Feb 2014
Accepted 3 Mar 2014
Available online 5 Apr 2014

Appendix 12: Published data on population structure of *P. falciparum* in Lagos and Ekiti States

Oyebola *et al. Malaria Journal* 2014, **13**:493
http://www.malariajournal.com/content/13/1/493

MALARIA JOURNAL

RESEARCH Open Access

Microsatellite markers reveal low levels of population sub-structuring of *Plasmodium falciparum* in southwestern Nigeria

Muyiwa K Oyebola[1,2], Emmanuel T Idowu[1], Haddy Nyang[3], Yetunde A Olukosi[2], Olubunmi A Otubanjo[1], Davis C Nwakanma[3], Samson T Awolola[2] and Alfred Amambua-Ngwa[3*]

Abstract

Background: Genetic diversity studies provide evidence of *Plasmodium falciparum* differentiation that could affect fitness and adaptation to drugs and target antigens for vaccine development. This study describes the genetic structure of *P. falciparum* populations in urban and rural sites from southwestern Nigeria.

Methodology: Ten neutral microsatellite loci were genotyped in 196 *P. falciparum* infections from three localities: Aramoko-Ekiti, a rural community; Lekki, an urban location and Badagry, a peri-urban border settlement. Analysis was performed on the genetic diversity, linkage disequilibrium, population structure and inter-population differentiation.

Results: Allelic diversity values were similar across all populations, with mean expected heterozygosity (H_E) values between 0.65 and 0.79. No matching multilocus haplotypes were found and analysis of multilocus LD showed no significant index of association. Genetic differentiation between populations was low ($\Phi PT = 0.017$).

Conclusion: The absence of detectable population structure of *P. falciparum* in southwestern Nigeria is evident in the lack of significant differentiation between populations separated by about 200 km. This implies that a fairly uniform malaria control strategy may be effective over a wide geographic range in this highly endemic region. However, more wide-scale survey across the country will be required to inform malaria control in this large and densely populated endemic region.

Keywords: Genetic diversity, *Plasmodium falciparum*, Linkage disequilibrium, Population structure, Genetic differentiation

Background

The incidence of malaria infections and malaria related mortality has reduced in many countries in Africa [1-3]. However, these successes remain limited in geographical coverage while transmission continues in some endemic regions in sub-Saharan Africa despite concerted efforts to reduce or eliminate the disease [4,5]. This is partly due to genetic diversity of the main agent *Plasmodium falciparum* which maintains population fitness against targeted interventions such as drugs [6,7]. Information on genetic diversity and parasite population trends that could help guide control programmes is lacking in regions with large human populations at risk such as Nigeria. The most recent report on patterns of malaria endemicity in Nigeria continues to show high levels of burden across the country with ~170 million people at risk [8]. This is despite more than a decade of vector control with insecticide-treated nets/long-lasting insecticidal nets (ITN/LLINs), indoor residual spraying (IRS), larval control and targeting of parasites with intermittent preventive treatment (IPT) and artemisinin-based combination therapy (ACT). With a proposed agenda for malaria elimination, it is important to determine the extent of genetic diversity, transmission intensity and the ultimate population structure of the parasites to support interventions.

There are various approaches to molecular determination of population structure including typing for polymorphic repeats in merozoite surface proteins (MSP 1

* Correspondence: angwa@mrc.gm
[3]Medical Research Council, Gambia Unit, Fajara, The Gambia
Full list of author information is available at the end of the article